Rutgers School of Communication and Information

media + the public interest initiative | news measures research project

Defining and Measuring Quality Journalism

This research is supported by the Democracy Fund and the Geraldine R. Dodge Foundation

March, 2015

Stephen Lacy
Michigan State University

Tom Rosenstiel
American Press Institute

Table of Contents

About the Authors

Stephen Lacy is a professor in the Michigan State University School of Journalism. He has written or co-written more than 100 refereed journal articles, more than 60 refereed conference papers, ten book chapters and four books, including a content analysis text, Analyzing Media Messages, which is in its third edition. He has co-edited two other books. He formerly co-edited the Journal of Media Economics. Since receiving his Ph.D. in journalism from the University of Texas at Austin in 1986, he has worked on multiple large-scale content analysis projects involving a wide range of media. The projects have been funded by the Knight Foundation, the Pew Foundation, and the National Science Foundation. His work has resulted in an MSU Distinguished Faculty Award and the Paul J. Deutschmann Award for Excellence in Research, a career award from the Association for Education in Journalism and Mass Communication (AEJMC). Professor Lacy served as director of the Michigan State University School of Journalism from 1998 to 2003, acting dean of the MSU College of Communication Arts and Sciences in 2014 and is a former president of AEJMC. Before becoming a professor, he edited a group of weekly suburban newspapers and worked as a reporter at a suburban daily, all near Dallas, Texas. He received his undergraduate degree in economics (1971) from the University of Illinois at Urbana.

Tom Rosenstiel is an author, journalist, researcher, media critic, and one of the nation's most recognized thinkers on the future of media. Before joining the American Press Institute in January 2013, he was founder and for 16 years director of the Project for Excellence in Journalism at the Pew Research Center in Washington, D.C., and co-founder and vice chair of the Committee of Concerned Journalists. During his journalism career he worked as media writer for the Los Angeles Times for a decade, chief congressional correspondent for Newsweek, press critic for MSNBC, business editor of the Peninsula Times Tribune, a reporter for Jack Anderson’s Washington Merry Go ‘Round column, and began his career at the Woodside

Country Almanac in his native northern California. He is the author of seven books, including The *Elements of Journalism: What News People Should Know and the Public Should Expect,* which has been translated into more than 25 languages and has been described as “The most important book on the relationship of journalism and democracy published in the last fifty years” (Roy Peter Clark, (Poynter), "a modern classic" (Bill Safire, *New York Times*), and one of the five "essential books" on journalism (*Wall Street Journal*). He and Kovach have also written two other books together, including, *Blur: How to Know What’s True in the Age of Information Overload.* His newest book is *The New Ethics of Journalism: Principles for the 21st Century,* co-edited with Kelly McBride of the Poynter Institute. His books and work at PEJ have generated more than 50,000 academic citations. He is a four-time winner of both the Sigma Delta Chi Award for Journalism Research and the national prize for media criticism from Penn State. Among his other awards are the Honor Medal for Distinguished Service in Journalism from the University of Missouri Journalism School, the Dewitt Carter Reddick Award for Outstanding Professional Achievement in the Field of Communications from the University of Texas at Austin, the Columbia Journalism School Distinguished Alumni Award and the Goldsmith Book Award from Harvard.

Introduction

The question of what constitutes quality journalism has taken on new urgency as the news industry undergoes its most profound disruption, probably since the beginning of commercial journalism. The news is no longer the province of a limited group of professional organizations that loosely agree on values or share common norms and training. New publishing companies, experimenting with different kinds of storytelling and revenue models, are reimagining what constitutes quality to capture audiences. Digital technology allows political actors, interest groups and advertisers to produce their own journalism that is designed to persuade as much as, or more than, to inform. The growing power of citizens, to choose what news they want and to act as producers as well as consumers, has elevated the importance of audience preference in defining quality. The growing influence of the audience has also given new momentum to the issue of teaching news and civic literacy to influence those tastes. In the age of low barriers to entering into publishing and high degrees of consumer choice, what is quality in journalism?

After reading the body of research and discussion about measuring journalism quality, it is obvious that some of the disagreement about how to do this, if it can be done at all, lies in the various definitions of the term "quality." This paper aims to come to grip with these disagreements, to identify points of consensus, and to offer some suggestions for the contemporary landscape based on that consensus. In so doing, we recognize that arguments about measuring quality run deeper than semantics.

While the issue of quality may be more pressing than ever, it is not new. In a 2004 essay, Leo Bogart (2004) discusses whether journalism quality can be measured in the context of newspapers and whether or not that matters. He closes by writing:

> Whatever the criteria they use, the conclusion is clear: a newspaper's investment in its news operation is likely to yield a solid return. What counts, however, may not be the dollar amount of that investment, but how it is spent—in short the quality. And how is that quality to be measured? (52)

The difficulty of answering that question stems from the complexity of defining journalism quality including the many ways it can be operationalized into measures and the multivariate process involved in transforming newsroom investment into quality. We agree with Bogart that resource investment alone does not determine journalism quality, although investment is an important element of achieving quality. Regardless of other factors, a newsroom with fifteen reporters will provide different content than a newsroom with ten reporters: journalists will have more time to do and review their work, the possibilities of beat expertise will rise, making it more likely to produce content that is better able to serve the community and individuals who use it. However, to test such a statement, one must be able to measure journalism content and its "quality." This paper will explore the current ability of journalism to do just that.

Getting Started

Examining quality journalism requires that we start with a definition of journalism, not as easy a task as it may sound. Disputes over what constitutes journalism have been around for as long as journalism has existed. The editors of the existing mercantile newspapers criticized the penny presses during the 1830s. Mainstream editors criticized populist and other alternative newspaper editors during the late 1900s as not being real journalists. More recently, professional journalists have criticized the citizen journalism movement. Concerns have also been raised about efforts such as sponsored content, as well as news being produced by organizations whose fundamental purpose is not producing news but selling their products or advancing a political agenda.

Rather than continue these arguments, this paper will start with a general definition of journalism and proceed to discuss the efforts that have been undertaken to measure the quality of that journalism.

> *Journalism is the serial presentation of information and conversation about public events, trends and issues distributed through various media with the primary purpose of informing, entertaining and connecting citizens in communities.*

This definition was derived from a review of the Report from the Commission on the Freedom of the Press (1947), also known as the Hutchins Commission Report, Kovach and Rosenstiel, (2001, 2007, 2014) and McBride and Rosenstiel (2013).

The level of evaluation and perspective of the evaluators are two fundamental dimensions in understanding disagreements about what constitutes quality journalism. Journalism can be evaluated at more than one level. Van Cuilenburg (2000) specified four levels—*content units* (program or article), *content bundles* (a channel or magazine), *medium type* (the bundles supplied by a medium to a market, such as television or print), and *society's communication system* (markets and communities). In addition, there is a level of multiple bundles from the same organization (several websites, apps, and other products from the same news organization). Moving upward from content unit to community results in an increase in the number and diversity of total content to be evaluated. All of these levels are affected by time because all of these levels could be measured a one point in time or across a series of years. A market could hold a diversity of voices at one time and then have that diversity disappear from market forces.

When discussing journalism and its quality, it is useful to specify the level. For example, source diversity is often cited as a goal for quality journalism (Napoli 1999, 2003), but the use of the term "source" will vary with the level. An individual–story may, for instance, not contain much diversity in sources, but diversity could exist among all the stories available that day about a given event or trend. Thus, the diversity available to a community would be greater if all the journalism products in the market are evaluated.

The same pattern of levels applies to journalism consumers. Not all consumers may be interested in diversity of viewpoints, but the U.S. Constitution is based on the idea that the community as a whole needs diversity. So standards of quality for journalism that provides utility to one person may not be the same standards used for evaluating utility provided to members of a large community. In other words, diversity can be evaluated from a variety of perspectives and at various levels of content and citizen aggregation.

Equally important is the recognition that groups of people who share particular information needs, political views, or have similar backgrounds may be more likely to agree on what constitutes quality than people with different responsibilities or perspectives. Teachers, for example, may differ in outlook from many parents or administrators in what they consider a “good” story about education. The members of different groups share a perspective that allows some level of inter-subjective agreement on what constitutes quality. However, an issue in establishing standards associated with a group’s perspective is that groups can be defined in many ways and the larger the group, the more likely individuals will vary from a common perspective. Some scholars argue that certain types of groups exist. Lacy and Simon (1993, 64) identified readers, journalists and critics as three legitimate group perspectives. Although generally true, these three are far from exhaustive and each group contains subgroups. For example, researchers have found areas of agreement and disagreement between editors at large dailies and small dailies as to what constitutes quality journalism (Gladney, 1990). The existence of up-scale and downscale publications in many countries also suggests variations among readers. As a result, one could generate a more extensive list of perspectives, such as journalists, journalism educators, academic scholars, high socioeconomic news consumers, low socioeconomic news consumers, non-academic professional critics and news organization managers. A story that is too detailed, long, or technical for one audience may strike another audience as highly informative and well sourced. Variations will exist within all the perspectives, and the amount of this variance is an empirical question that can be determined with research. The overall quality journalism standards that would be most useful are those that have high levels of inter-subjective agreement within perspectives, across-perspectives and across markets and cultures.

Is Quality Journalism an Important Issue?

The question of why we should be concerned with the quality of journalism begins with the question of why we should be concerned with journalism. One answer would be that journalism is the only business specifically protected by the U.S. Constitution. This protection reflects a fundamental assumption that individual decisions that aggregate to elect public officials are optimal when voters have access to large amounts of information and opinion. Journalism provides that information and opinion. It is a check against tyranny and abuse. More generally, journalism, by making information more transparent in society, is an essential ingredient for democratic self-government.

If one accepts the definition that quality represents the ability of journalism to fulfill its functions, then improving quality of journalism would improve the ability of citizens to use the journalism to make better decisions and provide a check against abuse and malfeasance by people in power. This notion may be most memorably embedded in the famous aphorism of Joseph Pulitzer: “our Republic and its press will rise or fall together.” The simple equation becomes—increasing quality of journalism will lead to better decisions by citizens and more accountability of government.

Why is Quality Journalism Measurement an Important Issue?

This rather simple equation, however, becomes more difficult when we try to measure journalism quality in order to study how it can be improved. In efforts to develop policy, journalism practices and education that would lead to higher levels of quality journalism can only be tested with reliable and valid measures of what constitutes quality. For example, if research indicates that market-level source diversity increases by increasing the number or journalism organizations, then foundations could invest in new journalism start-ups within a city. But establishing the relationship between source diversity and the number of journalism outlets requires a reliable and valid measurement of source diversity. Similarly, if research indicates that higher levels of investigative journalism correlate with more honest government, funders might want to increase the amount of investigative journalism. But we need ways to identify both investigative journalism and honest government. Today, it also may be important to discover if audiences are more likely to engage and learn civic information from one kind of presentation style over another. Doing so would require research on audience learning from different styles of news presentation.

What Will this White Paper Do?

In order to explore the concept of quality journalism, this paper will first review how quality has been defined and discussed by scholars and journalists. Second, it will examine previous efforts and methods used to measure journalism quality. Finally, the paper will provide discussion and questions for developing the next steps in creating measure for quality journalism.

Defining Journalism Quality

The meaning of the term *quality journalism* will vary from individual to individual, but because of socialization, meanings are more likely to be shared by group members who have some common experience. Two groups relevant to the literature about defining journalism quality are academics and professionals. Members of both groups share an interest in the nature and quality of journalistic output, but they have very different training and goals. (The group of professionals today also includes a growing number of data scientists working inside news organizations whose job it is to understand user engagement and behavior through digital analytics.) Professionals create journalism and academics study that journalism. The formal training they receive reflects those goals. These groups are not mutually exclusive. Some data scientists inside news companies have advanced research degrees. Some academics worked as journalists, and some journalists learn research methods. As a result, these groups have increasingly begun to interact and share their observations about quality journalism.

As an organizational device, the following section will divide the literature about what constitutes quality journalism into academic literature and professional literature, with the understanding that the distinction is somewhat arbitrary. The hope is that the two groups share some basic concepts of what constitutes journalism quality.

Academic

Academic researchers have examined journalism quality from the demand and the production side. The demand side emphasizes the interaction between the needs and wants of news consumers and the content. The supply side approach tends to specify characteristics of the content that are associated with higher quality levels. Both approaches typically define journalism quality as a matter of degree. It is not as simple as having or not having quality.

Demand Approach

In a model of news demand, Lacy (1989) proposed that news consumers evaluate content along several content attributes. The evaluation is based on how well those content attributes serve the information needs and wants of the individual evaluating the attributes. The news consumer's impression of quality of a journalism bundle is based on how well the content attributes in the bundle meet her or his needs and wants. For example, a person who lives in Grand Rapids, MI, and has a high interest in Big Ten basketball, the local school system activities, and Michigan government will not receive as much utility about those subjects from *The New York Times* as from the *Grand Rapids Press*. This doesn't mean the *Times* is void of quality. In fact, another person interested in fighting in the Middle East and the activities in the entertainment industry would perceive the *Times* as having high quality. That person may even be a Grand Rapids resident. The difference between these two quality evaluations lies in the difference between the individuals' needs and wants.

It is likely that at least some of the same journalistic elements (accuracy, completeness, fairness, etc.) would be found in the content of both news bundles. This is the difference between the demand and product approaches. The content could show the same elements of quality from a journalist perspective, but two different news consumers would evaluate the quality as being different because of their differences in information wants and needs.

The impact of quality from the demand perspective can take place at an individual or community level. The community level impact is more than just a summing of individual utilities because communities act as groups to govern themselves. In addition, there are individual information wants and needs that reflect the influence of membership in a social group. A person needs information to help decide how to vote not because she was born with the right to vote, but because she is a member of a social group that allows and may even expect individuals to participate in the political process.

To understand one level, it is important to know the process at the other level. It may well be that different levels require different "quality" attributes and measures. What has a positive impact for individuals might have a negative impact at the social level—for example, information that supports a person's biases may serve the individual but it may hurt the community because it does not allow the individual to adequately evaluate political candidates and make a well-informed voting decision.

An individual's motive for accessing and consuming journalism plays a role in his or her evaluation of quality. If a person, for example, would like to know about street repairs in her neighborhood, she might consider a story about the repairs to have high quality. Someone living in another neighborhood might consider the information useless. A research area that examines the motives behind media consumption is called "uses and gratifications" (Ruggiero, 2000). It has been criticized for being atheoretical and because the labels for various types of uses vary from study to study. Despite the variation in labels, Lacy (2000) identified four consistent types of individual media uses in the research literature (surveillance, social/cultural interaction, personal understanding and entertainment). He added decision-making as a fifth use. Surveillance involves keeping up with events in a person's surroundings. Social/cultural interaction involves any economic, political, cultural or social information about a social group. Personal understanding involves information that helps an individual understand herself. Entertainment information provides people with diversion and covers a range of activities. Finally, decision-making involves information that addresses specific decisions by an individual. This could involve buying a car, selecting a school system, voting in an election, or deciding what route to take on a trip. This information plays a specific purpose. High quality information in this instance allows a person to make an effective and efficient decision.

Just as there are information interests that reflect individual psychological needs, people have information needs and wants that reflect membership in social groups. At the risk of oversimplifying, these typically fall into five categories: *political/government*, information about governmental and political activities; *economic/business*, information about the operation and conditions of the economy and businesses within it; *social activities*, information about the broad range of events and activities among within a community, from sports and recreation to restaurants and more; *news you can use*, information aimed at helping citizens in a community address specific problems or understand defined phenomena that fall outside strictly civic, cultural or business realm, from questions about finding day care, to evaluating diets, to help with personal advice, parenting and more; and *cultural activities*, information about the artistic and cultural output within a community. These social needs and wants manifest themselves through the actions of individuals, but the demand for the information comes from the interactions with others.

At the Project for Excellence in Journalism and the American Press Institute, Rosenstiel has developed different batteries of questions that have probed different ways people use news. Initially, the grid involved five broad categories of news use (Purcell, et al., 2010):

- As a form of social flow, or so people could talk about news with family, friends and colleagues
- To be informed as a civic responsibility or to be a better citizen
- To find information that will help them with their lives
- Because they find the news relaxing, diverting or entertaining

- To help them professionally

More recently, Rosenstiel and his colleagues (2015) have made that list more specific, though some of the reasons can still be grouped into larger lists. The newer list being used in new surveys includes, in addition to those above:

- To help people decide where they stand on issues
- To help them feel more connected to community (more passive or psychological than actually talking about the news)
- To help people decide places to go or things to do (a subset of helping their lives)
- To help them take action on issues they care about (connected to citizenship)
- To stay healthy (a subset of helping their lives)
- To solve problems (subset of helping their lives)
- To save money (subset of helping their lives)
- To help raise their family (another subset of helping their lives)

When asked in various surveys, the answers usually come back the same way. Citizenship and talking about the news are always the top reasons (sometimes trading places). Helping people decide where they stand on issues is also a major reason, one connected to citizenship. The reasons that have to do with taking action, whether it is on issues that they care about, helping them solve problems, spend time, save money or help their careers all come lower on the list.

Overall, the demand approach toward journalism quality aggregates individual consumer's perceptions of how well journalism serves their needs and wants. This relativistic approach helps explain the arguments that quality cannot be measured because it is subjective. This argument assumes the needs and wants of individuals share little overlap, which would result in a huge range of opinions about what constitutes quality journalism. The assumption of such a range runs counter to the foundations of psychology and sociology, which are based on the assumption that humans have patterns of behaviors and motivations based on fundamental physical and mental needs.

Product Approach

In contrast to the demand approach, the product approach assumes there are characteristics inherent both in the messages and also in the bundles of messages that can be altered to improve the quality of the content. Journalists, managers, critics and scholars identify the characteristics they believe represent quality. One advantage of the product approach is that journalists and publishers can control or influence the character of the content they produce more easily than they can infer the interests of audiences. With a focus on quality within the product, producers are able to improve or do something about quality.

The product approach to journalism quality preceded the demand approach in academic research. Journalism ethicist John Merrill and colleagues discussed quality of newspaper organizations over a number of years (Merrill, 1968; Merrill & Lowenstein, 1971). In identifying how to measure quality, they distinguished between internal factors, which the news organization could control (e.g., staff quality), and external factors, such as how often they influence policy makers. In 1983, Merrill and Oden (1983, 110-111) identified a mix of reputational, content, and organizational elements, which could be measured in various ways.

McQuail (1992) approached the idea of media quality with the concept of *media performance*, which he defined in a way that was largely limited to public interest criteria, relating more to government and political news:

> "The independent assessment of mass media provision according to alternative 'public interest' criteria, by way of objective and systematic method of research, taking account of other relevant evidence and the normal operating conditions and requirements of the media concerned." (McQuail, 1992: 17, as cited in del la Piscina, 2014)

McQuail (2005) later elaborated on the idea of media quality and its benefits, identifying five basic values: freedom, equality, diversity, truth and information quality, and social order and solidarity. He also asserted that content determines the performance of a media outlet or system with regard to some of these values.

The concept of *freedom* refers to the ability to express oneself and to publish messages without being constrained by governments and economic systems. McQuail (2005) argues for a notion of performance that reflects the freedom value that is difficult to define and measure; however, he suggests that media outlets express freedom through an active and critical editorial policy and by not conforming to norms in the system. Freedom is a journalistic value at the systemic level. It exists to some degree in many countries, but the variance happens at the societal/community level and less at the organization level. The ability to be free does not guarantee the use of the freedom.

A second value, *media equity*, assumes that access to media should be open and equivalent for varying points of view and organizations and that those points of view be proportionate to some criteria. In other words, media access should not just be for the wealthy and elite social groups. The idea of access is a fundamental assumption for the efficient and effective working of the marketplace of ideas. Access occurs at a systemic level and is closely related to freedom and to diversity.

For McQuail, diversity as a value occurs at both the social and outlet levels. It assumes that a variety of viewpoints are available from a variety of sources. In discussing how performance might be measured at the content level, he suggested several dimensions (2005: 198): “genre, style of format in culture or entertainment; news and information topics covered; political viewpoints; and so on.” McQuail’s diversity concept as a value applies at the individual message level and the content bundle level. It is related to the idea of balance and fairness, which has been used as a quality measure (Simon, Fico, & Lacy 1989).

Truth and information quality is McQuail's fourth value. He used the term "objectivity," but divides it into factuality (truth, relevance and informativeness) and impartiality (balance and neutrality). He summarized the value by writing:

"The main information quality requirements are as follows:

- Mass media should provide a comprehensive supply of relevant news and background information about events in the society and the world around.
- Information should be objective in the sense of being accurate, honest, sufficiently complete and true to reality, and reliable in the sense of being checkable and separating fact from opinion.
- Information should be balanced and fair (impartial), reporting alternative perspectives and interpretations in a non-sensational, unbiased way, as far as possible." (McQuail, 2005: 202-203)

McQuail acknowledged the limits of the concept of objectivity, which often results in being bogged down in a philosophical discussion that is basically a disagreement on the underlying assumptions about human behavior and the role of science. He presents a discussion of these limits (2005: 355-358) and suggests that the presence of some bias does not mean content cannot adequately reflect elements of everyday behavior that would serve useful.

The final value listed by McQuail connected with information quality is social order and solidarity. This value relates to the role of media to assist society in creating "harmony" and "integration." This value need not assume that harmony equals conformity or order equals control. Media can promote pluralism just as it can support a dominant point of view. This value is a system-level value that is related to diversity when applied at the message or content bundle levels.

Of McQuail's five values, the two with the most relevance to evaluating the quality of journalism products (messages, journalism bundles and the output of journalistic organizations) are "diversity" and "truth and information quality." His discussion of these two values is reminiscent of the discussion of journalism by the Commission on the Freedom of the Press (1947), also called the Hutchins Commission. The Commission report stated that the press, which included radio and print at the time, should meet five requirements:

1. A truthful, comprehensive, and intelligent account of the day's events in a context which gives them meaning;
2. A forum for the exchange of comment and criticism;
3. The projection of a representative picture of the constituent groups in society;
4. The presentation and clarifications of the goals and values of the society;
5. Full access to the day's intelligence. (1947, xi).

As a whole, the Hutchins Commission report codified the idea of journalism as responsible to its citizens. It further drew the customary distinction between news reporting and opinion (requirement 2). Requirements 3, 4 and 5 reflect the need for diversity in journalism. Requirement 1 and 5 state that journalism should be founded on truth and be written in a way that is accessible, comprehensive, complete, and contextual.

One consistent element in much of the academic literature on journalism product quality is the concept of diversity. Diversity connects the demand and product approaches. A diversity of sources is more likely to meet the needs and wants of a wider variety of consumers than journalism with a limited number of sources that reflect a partial set of perspectives. An article, for example, that examines the impact of the Patient Protection and Affordable Care Act on a wide range of socioeconomic groups would serve a greater range of news consumers than one that dealt only with the impact on families whose income is below the poverty line.

Napoli (1999, 2003) pointed out that federal policy has traditionally aimed to maintain diversity in the marketplace or ideas, which he calls the content market. He differentiated among sources diversity, content diversity, and exposure diversity (1999). *Source diversity* concerns ownership, programming and the workforce and addresses the number and nature of journalists and journalism outlets providing news and opinion. *Content diversity* is divided into program-type format, demographic, and idea/viewpoint. *Exposure diversity* addresses the consumer's role in accessing diverse content and is divided into horizontal and vertical. *Horizontal diversity* concerns the diversity across all channels available to consumers, and *vertical diversity* concerns the diversity within individual channels. A diverse vertical channel would be a news organization that contains a great deal of content diversity in its products. Horizontal diversity would require that the sum of outlets available in a market contain a great deal of diversity. The concept of vertical and horizontal diversity is consistent with the classification of content levels (messages, product bundles, product bundles across time, and markets).

Policy tends to deal with the market level of content diversity as a phenomenon that is assumed to develop from channel (journalism organization or delivery channel) source diversity. Napoli (1999) discussed this connection, but there is a more fundamental one, which is that channel diversity (whether a given product bundle or all the bundles produced and delivered through a channel) reflects the diversity within the individual stories. Because news reporting, unlike opinion, depends primarily on receiving information from sources (people, documents, etc.), if individual messages are not diverse, the probability of having a diverse bundle of messages declines. In other words, there is a positive correlation between diversity found in individual messages and found in the bundles of messages that they form.

Interest among academic researchers about measuring quality journalism is not limited to the United States. In a collection of international case studies, Anderson, Ogola, and Williams (2014) discuss how quality hard news can be provided given the current upheaval in the news industry. They start with a recognition that quality is multifaceted and that they must narrow the range of their discussion and measurement. They argue for a framework for analyzing quality based on five Cs and one A—

comprehensibility, context, causality, comparativeness, comprehensiveness, and accuracy. These are generally consistent with the quality elements presented elsewhere. Most of these terms are self-evident, but two could use more detail. The term comparativeness involves presenting more than one side of an issue and is related to diversity, fairness, and balance. The term causality calls for an explanation of the most likely causal factors regarding the events and issues being covered. This is an element that is less often found in quality classification systems, perhaps because of the difficulty of identifying the "true" causal relationships.

In a study of the democratic quality of radio news in Liberia, Spurk, et. al. (2010) used a classification system with four main quality elements and then sub-elements. These were: 1. *Information* (comprehensiveness, diversity of topics, diversity of actors in the stories, diversity of sources), 2. *Orientation* (presence of background information, presence of reasons, causes beyond factual information, diversity of viewpoints, clear separation of fact and opinion-transparency), 3. *Forum* (diversity of viewpoints, diversity of sources, balance of opinion), 4. *Scrutiny* (presence and balance of critical viewpoints, own journalistic inquiry). Researchers applied these categories to 246 news stories from four Liberian stations in 2007. Although two coders coded each story, there was no reliability coefficient reported for the results.

Both the product and demand approaches toward journalism quality aim to identify specific characteristics of news content. They differ on how those characteristics are evaluated. The product approach identifies a specific set of quality characteristics that can be applied to all journalism content. The demand approach, however, is more relativistic. It evaluates content on how well it meets the psychological needs and wants of specific individuals. Theoretically, the individuals' utility could be aggregated if it could be measured. The product approach also delineates characteristics that meet community-level functions, such as information about political, cultural, and social activities.

Professional

The discussion and definition of quality journalism by professionals has tended to take a product approach, usually listing elements of quality. Again, this may be because it was easier for producers of journalism to focus on what they could control, the content.

Working with the Project for Excellence in Journalism, as an example, Rosenstiel, et al. (2007), developed a comprehensive scheme for assessing and assigning quality scores to local television newscasts in the late 1990s. The work was later described in the book, *We Interrupt This Newscast*.

The group assembled a team of highly respected local television news professionals to identify the elements of a quality newscast. These professionals first completed a detailed questionnaire and then they were assembled in a day-long discussion to more fully understand their thinking. The professionals agreed that newscasts should, a) cover the whole community; b) cover significant and substantive issues; c) demonstrate enterprise and courage; d) be fair, balanced and accurate; e) be authoritative; f) be clear and not sensationalize. The team then developed a coding scheme that reflected those values and tested whether viewers recognized them by showing high and low scoring stories to focus groups. In addition, the study examined whether these variables (when quantified in content analysis) correlated to better long-term ratings trends.

A. To assess that a newscast covered the whole community, the scheme identified the ***primary topic of each story*** and scored newscasts higher that had a higher breadth of topics over the course of a week.

B. To assess that the newscast was covering substantive and significant issues, each story was evaluated for ***focus***. The focus of a story examined the degree to which a story touched on underlying themes, ideas, trends or issues raised by the incident or put the news in a larger context. A story about a car accident was at one level for focus. A story that touched on how many accidents had occurred at that intersection and probed why would be higher. There were 10 levels of focus in all, with stories about ideas higher than those about celebrity gossip.

C. To assess courage and enterprise, the coding scheme measured ***the level of effort*** that went into the story, from a multi-part expose, to a pre-arranged event such as a press conference, to footage shot without a reporter being on scene.

D. To measure accuracy and fairness, the coding scheme evaluated each story by two criteria—the ***number of sources*** cited in a story and how many ***differing viewpoints*** were contained in any story that involved a disagreement or inherent controversy. Stories presented as undisputed (a fire happened) were noted separately.

E. To measure authoritativeness, the coding scheme evaluated the level of ***expertise of each source*** cited in a story. Expertise varied by topic. A qualified brain surgeon would be a credentialed expert on a story about brain surgery. A citizen on the street would be a qualified expert, however, in responding to the president's state of the union.

F. Clarity and sensationalism were measured fairly simply. If a story was so attenuated it made no sense, it lost a point for clarity. Sensationalism was defined as a piece of video being repeated beyond the point it conveyed any new information, and being replayed simply to excite or exploit.

It is notable how similar these values – identified by local TV news professionals out of a mix of professional mission as well as an experienced sense of what attracts viewers – are to the values developed more theoretically by academics. Elements such as source, topic, and viewpoint diversity; contextualizing or orienting news; watchdog reporting or holding the powerful accountable; and various other themes appear on both lists.

There have been other attempts to quantify product quality. Leo Bogart, who worked for many years at the Newspaper Advertising Bureau (1989), surveyed 746 daily newspaper editors and came up with a rank ordering of 23 characteristics of high quality newspapers. These are:

1. *Ratio of staff-written copy to wire service and feature copy*
2. *Total amount of non-advertising content*
3. *Ratio of news interpretation and backgrounders to spotty news reports*
4. *Number of letters to editor per issue*
5. *Diversity of political columnists*

6. *High "readability" score*
7. *Ratio of illustrations to text*
8. *Ratio of non-advertising content to advertising content*
9. *Ratio of news to features*
10. *Number of staff-bylined features*
11. *Ratio of sports news and features to total news content*
12. *Presence of news summary*
13. *Presence of an "action line" column*
14. *Number of editorials per issue*
15. *Number of wire services carried*
16. *Ratio of cultural news, reviews and features to total news coverage*
17. *Ratio of service journalism news to total news coverage*
18. *Ratio of business news, features to total news coverage*
19. *Number of political columnists*
20. *Number of comic strips*
21. *Length of average front page story*
22. *Presence of an astrology column*
23. *Ratio of state, national, world news to local news*

Most, but not all, of Bogart's characteristics are also consistent with the discussion by academic researchers. A few (e.g., presence of an astrology column and number of comic strips) seem to be from another era. Currently, both types of content are readily available online. In addition, wire services are easily available directly from the services via mobile and Internet distribution.

Bogart's survey is important because it has been replicated and supported (Meyer & Kim, 2003) and operationalized through content analysis (Lacy & Fico, 1990, 1991; and St. Cyr, Lacy & Guzman, 2005). Meyer and Kim conducted a survey of 285 editors from the American Society of Newspaper

Editors in 2002. They listed 15 quality characteristics, including 13 of the top characteristics from Bogart (1989). The Spearman's correlation between the rankings of common characteristics in the Bogart and the Meyer and Kim surveys equaled .77, which shows consistency across time. The top five characteristics in the Meyer and Kim survey were: 1. vigor of editorials, 2. total amount of non-advertising content, 3. high ratio of staff-written copy to wire service and feature service copy, 4. diversity of political columnists, and 5. number of staff-bylined stories. Three of these were in the top five of Bogart's survey. Number 1 was not a characteristic in the original survey and number 5 ranked 10th in the original survey.

Another typology for quality journalism was suggested by Denton (1999) and adopted by Thorson, et. al. (2008). The quality attributes included proximity, safety, utility, government, education, spirituality, support from the community, identification with others and the community, recognition for achievements of self and others, empowerment, and credibility.

The Committee of Concerned Journalists, a group co-organized by Rosenstiel and fellow journalist Bill Kovach, conducted extensive research on journalists in the late 1990s to assess whether any broad consensus existed over values among professionals working across different media and journalism disciplines. The Committee held twenty-one public meetings across the country, each one gathering testimony on different issues about different values from prominent journalists and members of the public. Surveys were also conducted with journalists over values in order to test the conclusions arrived at during the public sessions. That work was distilled in the book, *The Elements of Journalism: What Journalists Should Know and the Public Should Expec*t (Kovach & Rosenstiel, 2001, 2007, 2014), which has become a popular text in journalism schools. *Elements* identified first that:

> *The primary purpose of journalism is to provide citizens with the information they need to be free and self-governing.*

This was meant in the broadest possible terms to include news about culture, sports, popular culture, society, and more.

To fulfill this purpose, the group identified ten principles or elements that journalists should adhere to in order to approach quality work:

1. *Journalism's first obligation is to the truth.*
2. *Its first loyalty is to citizens.*
3. *Its essence is a discipline of verification.*
4. *Its practitioners must maintain an independence from those they cover.*
5. *It must serve as a monitor of power.*
6. *It must provide a forum for public criticism and compromise.*
7. *It must strive to make the significant interesting and relevant.*
8. *It must present the news in a way that is comprehensive and proportional.*
9. *Its practitioners have an obligation to exercise their personal conscience.*
10. *Citizens have rights and responsibilities when it comes to the news as well—even more so as they become producers and editors themselves.*

Those principles, which drew consciously on work of the Hutchins Commission and others, have proven resonant with students and educators since their initial publication in 2001.

More recently, the question of audience engagement has become a more central concern in thinking about quality. This has largely reflects the increased ease in the digital age to measure audience preferences with metrics about reading time, visits, and sharing data that did not exist before. Perhaps just as important, the news itself has been atomized. Rather than engagement with a news organization being measured in the form of ratings to a newscast or circulation of a publication quarterly or annually, news professionals now must consider the engagement with individual pieces of content and contend with what values they want to put on that content. What matters more: The number of clicks to a given story? Or the depth of engagement on that story? Are there ways to consider engagement with issues, not just individual stories? Is the rate at which users share a story a sign of its higher value to readers (something that the publication Buzzfeed is focused on through its concept "social lift," because it signals that the reader not only consumed it but wanted to associate him or herself with it)? Or is time spent with a story more important (a metric that the rating service Chart Beat and the publishing platform Medium focus on). Or is it even possible to measure whether a diversity of topics (through the right mix) can broaden the audience of a publication in a way that targeting only popular stories cannot?

The American Press Institute has taken steps to try to understand the relationship between quality and these different metrics of engagement with an initiative and an app it calls "Metrics For News." Using the app, publishers meta tag their journalism content for various journalism qualities. They then can correlate those journalism characteristics to whatever digital analytics those publishers have. The program lets publishers understand a variety of concepts they cannot grasp from conventional analytics alone—such as the relationship between higher levels of enterprise and engagement, accountability journalism and engagement or what kind of journalistic work does better within each coverage area.

While the list of meta tags can be customized by publisher, among the elements the program can tag include:

- What as the primary topic of the story?
- What level of effort or enterprise went into it (a hierarchy from wire brief to expose)?
- Who was the author (reporter, freelancer, user generated, etc.)?
- Whose idea triggered the story?
- What is the fundamental value of the story to the use (what job or jobs does it do)?
- Story length
- Style of presentation (analytic versus neutral and multi media format)

When these characteristics are correlated to conventional digital analytics, such as page views and time, the publisher's data are transformed into journalism metrics. The program allows publishers to see what characteristics of coverage of a given topic drive more page views, more sharing, more time per story and other metrics.

The program also goes one step further in assessing quality. It can combine the various analytics available about reader engagement into a customized single number, or analytics index, which can combine page views, time, loyal users versus casual ones, local readers versus out of market, and more. The key concept behind the idea of indexing as it pertains to quality is that it forces a publication, or even a company, to decide what kind of engagement matters to them. Are they more interested in deeper engagement (time per story)? How much do repeat visits matter (loyalty)? How important is local readership versus all readership? How do these different valuations balance or blend? Indexing forces them to decide these issues, and to even settle on quantitative values for them.

Summary

The discussion of journalism quality by the various scholars and professionals takes three approaches. One is to specify the characteristics of quality-oriented news organizations. The second is to specify the content attributes that reflect the output of these organizations. The third is to analyze engagement data to see what kind of qualities resonate with audiences. Distilling observations in these efforts, here are some of the common content quality characteristics that have been suggested:

1. *Presentation quality* – High quality journalism should have high production quality and be accessible to a wide-range of community members.
2. *Trustworthiness* – High quality journalism should be accurate and credible, which can be measured in various ways.
3. *Diversity* – High quality journalism should be diverse in the number of sources, issues, topics and ideas within information bundles and across the community.
4. *Depth and breadth of information* – High quality journalism should provide a deep and broad understanding of issues and trends. In addition to looking in depth at important issues, this involves efforts to provide context through both news reporting and opinion. This could be

said to include, according to the API data, a range of story presentation styles, as that correlates to broader and deeper audience engagement.

5. *Comprehensive* – The bundle of quality journalism available in a community should cover the range of important community events, issues, and trends. This is measured at the community level, whereas depth and breadth are measures of range at the publication and story level.
6. *Public affairs* – Bundles of information from outlets should include a strong emphasis on public affairs reporting and opinion about government, politics, schools and other activities that affect the quality of life in the community.
7. *Geographic relevance* – Bundles of journalism should devote a significant amount of their content to events, issues and trends of importance to people in their primary geographic coverage areas.

These characteristics are applicable at multiple levels and from multiple perspectives. Some of the characteristics flow from common types of organizational characteristics and newsroom investment. For example, a large, well-trained staff could contribute to presentation quality, comprehensiveness, diversity, and depth and breadth of information.

Existing Efforts to Measure Quality Journalism

Journalism quality has been measured using a variety of methods from both the demand and the product side. Held and Ruβ-Mohl (2005, 56ff) identified four ways that quality has been measured, as reported by Gertler (2013, 15):

1. "Direct measuring of quality criteria, for example through content analysis; dimensions would be topically, interactivity, transparency, comprehensibility, relevancy right through to objectivity.
2. Audience preferences, because the audience has its own idea of quality; here, market research is deployed.

3. Expert judgments that can be found in sanctioning environments (Press Council) or also in juries; scientifically supported by multi-stage Delphi Method.
4. Indirect indicators that do not measure the product, but rather the production conditions; for example the number of editors, educational level and work experience, budget, etc."

This classification system will be used to examine how researchers have actually measured quality journalism. Some of the following studies include more than one approach.

Content Analysis

Content analysis is a commonly used method in journalism research. In can be used just to describe media content or to examine the relationship between content and its effects, and between content and the antecedent variables (e.g., competition, ownership, newsroom routines) that affect its creation (Riffe, Lacy & Fico, 2014). This review of research will concentrate on content analysis studies that explicitly address content quality or are somehow related to the general discussion of elements that give journalism higher levels of quality.

Fico and his colleagues (Simon, Fico, & Lacy, 1989; Fico, Freedman, & Love, 2006; Fico & Freedman, 2008; Zeldes, et. al., 2008) have published a series of studies that used content analysis to assess the balance of political news coverage. Balance within and across news organizations is necessary to provide diversity. They studied two types of balance: *partisan balance*, which measures space given to sources with a particular political orientation (conservative or liberal), and structural balance, which measures systematic imbalance (space allocated to candidates) regardless of the favored candidate's political orientation. The partisan balance index identified the partisan sources favoring a candidate and calculated four measures: "(1) Whether partisans for one candidate had more paragraphs of assertions than the opponent's partisans; (2) whether partisans for one candidate made exclusive assertions in the lead; (3) whether partisans for one candidate made exclusive assertions in the second through fifth paragraphs; (4) whether partisans for one candidate made exclusive assertions in the sixth through tenth paragraph" (Fico & Freedman, 2008, 505). Each component was evaluated as supporting the conservative (Republican) or liberal (Democratic, Independent or Socialist) candidate. The index could range from -4 favoring one candidate (e.g., conservative) to +4 favoring the other (e.g., liberal).

Using the indices, Fico, Freedman, and Love (2006) studied 9 newspapers' coverage of 2004 U.S. Senate races and concluded the treatment was fairly even, but news processing routines rather than partisanship produced the imbalance that was found. Fico and Freedman (2008) studied coverage of the 2006 Senate election in eleven newspapers and concluded most stories were biased toward Democratic and other liberal candidates. They also found newsrooms with a higher proportion of women were more likely to have balanced coverage. Zeldes, Fico, Carpenter, and Diddi (2008) used the same indices to study television coverage of the 2004 presidential election. They found that broadcast and cable news provided more attention, time and prominence to John Kerry than to George Bush, but the broadcast networks were more balanced in their attention than were the cable channels.

These measures tap into the allocation of sources and space among candidates in political news coverage. However, the indices do not directly address the biased wording that can be found in stories. In other words, the indices treat all biased sources as if they have equivalent impact on readers. A test of validity would be to use the indices to independently assess the degree of bias in the language and compare the two assessments.

The Committee of Concerned Journalists worked closely with The Project for Excellence in Journalism and, in addition to its comprehensive attempts to create quality scores for local television, examined other dimensions of quality across various media. These included such issues as how stories were framed (more comprehensive framing was better.) If all campaign stories, for instance, focus on tactics, citizens may never learn much about issues. Other studies tried to measure political bias by looking at the viewpoints in stories and identifying whether the preponderance (usually a using a two-thirds rule) of quoted sources reflected the views of one political party over another's. If a publication consistently reflected the views predominantly of one party it was said to have a clear political orientation, even if individually stories might appear to be well sourced and even contain some soft measures of balance and depth.

Rosenstiel at PEJ also extended one of the elements from the organization's local TV news studies into a broader context. From 2007 until 2012, PEJ analyzed major outlets from a range of media to see the breadth of topics covered, in what it called the News Coverage Index. That work identified such trends as the narrowing of focus of cable news as it concentrated on a few big stories and on politics, and the fact that the evening network newscasts were the last place on television where one might find a national news agenda as broad as one seen in newspapers.

A similar approach to topic diversity has been employed by others. In a study of quality in 350 German newspapers between 1989 and 1994, Schönbach (2000) reported on a 1997 study (Schönbach, 1997) that measured quality by the variety of topics across newspaper sections and within the sections, the amount of background information and commentary, the amount of coping news, emphasis on local news, the amount of soft and entertainment news, the efforts to use design to structure the newspaper, use of visual devices to present information, legibility of type, the overall appearance of the paper, attempts to engage readers, and responses to a survey about marketing. He found strong relationships between content variation and circulation. He concluded:

> Instead, serious information paid off for local dailies. That does not mean that it had to be sophisticated or boring. But readers seemingly did not want to read about gossip in politics or economics sections. This result applied to all criteria for success, target groups, and types of newspaper competition. (2000, 94)

Another approach that explicitly aimed to measure newspaper journalism quality adapted the Bogart (1989) survey of editors to a content analysis. Lacy and Fico (1990, 1991) used variables from a content analysis by Lacy (1986) to operationalize the Bogart survey results. The study used seven of Bogart's measures with some modifications. These were high ratio of staff-written copy to wire service copy; total amount of non-advertising copy; high ratio of news interpretations and backgrounders to spot news reports; high ratio of illustrations to text; number of wire services carried; length of average front page news story; and high ratio of non-advertising content to advertising. In addition, the index included the square inches of copy divided by the number of reporters listed with bylines, which was assumed to be a measure of reporter workload. The content analysis was conducted on a randomly constructed week from November 1984 for a randomly stratified sample of 80 U.S. daily newspapers in one study (Lacy & Fico, 1990) and 114 U.S. dailies in another (Lacy & Fico, 1991).

Because the eight measures used various scales (e.g., square inches, number of wire services, ratios of inches), each of the measures was converted to a z-score and then summed to create the index. The initial study (Lacy & Fico, 1990) adjusted the index for circulation size and then ranked 15 newspaper groups. The top ranked group was the New York Times. The 15th ranked group was Clay, and 14th ranked group was Gannett. In the second study (Fico & Lacy, 1991), the quality index for 1984 was used to predict circulation in 1985, controlling for city population. About 22% of the variation in 1985 circulation was positively and uniquely related to 1984 quality.

In a follow-up study, St. Cyr, Lacy, and Guzman-Oretga (2005) used the 1984 quality index to predict circulation in 1990 and 1995. The goal was to see if the relationship between quality and circulation found in the 1991 study (Lacy & Fico) would maintain itself across time. Controlling for number of households in the county, the regression found that about 23% of the variation in 1990 circulation was related uniquely with the 1984 quality index. By 1995, the relationship was not statistically significant and had dropped to 9.5% of variance.

The quality index used in these studies suggests that characteristics discovered in surveys of professionals can be operationalized with content analysis. One drawback is that the index had to be standardized because of incompatible measurement scales. Standardization requires large sample sizes to be effective. A possibility solution would be to develop a quality measures with a simpler scale that was still based on a survey of journalists.

A number of content analyses have examined geographic topics, with an emphasis on local coverage as a measure of quality. For example, Lacy and Sohn (1990) measured the square inches of news in metro dailies and suburban weeklies devoted to specific suburbs around Detroit and Denver. They used a constructed week from February 1986. The types of local content included display advertising, insert advertising, local government, local law enforcement, local sports, local social news, and local business, all measured for each suburb. The authors also obtained the circulation of weeklies within given suburbs and the circulation of the four dailies (two in Detroit and two in Denver). The square inches of content about a given suburb correlated positively with daily and weekly circulation in that suburb, although the degree of correlation for a given topic varied with type of newspaper and metro area. Multiple scholars also content analyzed local television station news for local emphasis (Adams, 1980; Wulfemeyer, 1982; Bernstein, et. al. 1990).

Some studies have used content analysis to study context and in-depth reporting. Zeldes, Fico, and Lacy (2008) examined broadcast television coverage of the 2004 Democratic presidential primaries for reporter speculation, multiple viewpoints in a story, story emphasis, anonymous sourcing and source transparency. They compared the primary coverage to all other stories and found primary coverage was more likely to include reporter speculation, more likely to fully identify sources and more focused on winners and losers than non-primary sources.

Bernstein and Lacy (1992) coded stories from fourteen local television news stations for amount of coverage of government and in-depth coverage of local government. In-depth coverage involved two stories presented as a package about the same issue involving local government. About eleven-percent of the 286 local government stories at large-market stations and three-percent of the 108 local government stories at the small-market stations were classified as in-depth.

Local government was the subject of a content analysis of a national sample of daily newspapers, weekly newspapers, citizen journalism sites, local radio stations and local television stations conducted by a team of researchers at Michigan State University (Fico, et. al., 2013a; Fico, et al., 2013b, Fico, et. al., 2012; Lacy, et. al., 2013; Lacy, et. al., 2012). The project, which studied a random sample of more than 800 news outlets from throughout the United States, not only examined the amount of city and county government coverage but the sourcing as well. This included both the number of sources in the articles and a source diversity index. The source diversity index was relatively simple. Each article was coded for the presence or absence of nine types of human sources (government officials, business sources, etc.) and seven types of written sources (government meeting minutes, reports, etc.). If one or more of the source types was present in the story, the article received a one. The articles with ones were summed to calculate the value of the index. The index could run from zero (none of the types) to sixteen of all of the types. Daily newspapers had the higher diversity index, but the average was only 1.48 for government meeting stories and 1.62 for non-meeting stories (Lacy, et. al., 2012).

In a recent study of four New Zealand daily newspapers, Gibbons (2014) measured content in a variety of ways: percentage of pages devoted to news, percentage of pages devoted to advertising, percentage of news written by the newspaper's staff, amount of space devoted to various topics, and amount of space devoted to various geographic locations. As a measure of quality, he used the pages of news divided by the logged local population, which controlled for the more diverse and extensive information needs of larger communities. Even though the sample included only four dailies, he found a weak positive correlation between quality and readership. This is consistent with studies in the United States that have found relationships between content quality and circulation (Lacy & Fico, 1991; Lacy & Sohn, 1989; St. Cyr & Lacy, et. al., 2005; Cho, et. al., 2004; Stone, Stone & Trotter, 1981).

Del la Piscina, et. al. (2014) based their analysis of five European dailies over 12 years on McQuail's (1992) discussion of quality. Their quality index included three elements (252-253): *selection process* ("a quote from the source of the information, the nature of the sources, the factual aspect of the matter, degree of topicality and newsworthiness); *preparation process* ("accuracy, depth, presence of different perspectives within the item, contribution made by other information elements, and correctness of journalistic language); and *social contribution* ("power watchdog, promotion of social debate, respect for human dignity, presence of cultural references from other countries, and the combating of social marginalization"). Each area was given points—2.5 for selection process, 5 for preparation process, and 2.5 for social contribution. The points were totaled for a scale of zero to 10. The study does not really fit the content analysis method because no reliability coefficient was reported and the manuscript lacks detail about how the quality classification system was applied (see Riffe, Lacy & Fico, 2014).

Some articles have suggested variables to use in a quality content analysis, but they contain no actual application of these variables to content. For example, Gertler (2012) discussed quality indicators, which he said can be seen in the output of the journalistic process. Diakopoulis and Essa (2008) suggested an annotation model for online video. They mentioned a specific annotation system called the Videolyzer, which allows typed comments about the presence or absence of quality indicators as a coder watches video. They provided six key quality indicators: validity, accuracy, reliability, bias, transparency and comprehensiveness (fairness and thoroughness). They provided a decision tree that can be used in the application of the quality indicator. The article provided no discussion of establishing reliability of the annotation process.

Another area of research about quality studies content by applying standards of grammar, punctuation, and accuracy. For example, Meyer and Arant (1992) used an electronic database from fifty-eight dailies to evaluate two style errors, one spelling error and one grammar error. These were picked because of the common misuse and were: the use of "most unique," the Associated Press (AP) preferred spelling of "minuscule," using of the AP preferred spelling of "judgment," and the presence of the misspelling of "accommodate." The study found positive, statistically significant correlations among the four measures, which suggests that the selected type of error tapped into an underlying process. They also translated the errors into z-scores and summed them. This gave an editing quality index that ranged from .92 at the *Akron Beacon Journals* to -2.64 for the *Annapolis Capital*.

As a test of validity, Meyer and Arant (1992) correlated the editing index with the number of Pulitzer Prizes and newspaper circulation. The relationship between the index and number of Pulitzers was curvilinear. After transforming the Pulitzer awards to the square root, the correlation between editing and number of prizes equaled .335, which was significant at the $p = .01$ level. The correlation between the error index and circulation equaled .311 and between number of Pulitzers and circulation equaled .546.

Another research area related to editing concerns accuracy. Work in the area of accuracy goes back to the 1930s with Charnley (1936) and continues today. The results of studies have consistently found that accuracy falls into objective inaccuracies, which involve errors that are verifiable (names, locations, time), and subjective inaccuracies, which concern differences between reporters' and sources' interpretations of source statements.

Maier (2005) surveyed 4,800 news sources to assess accuracy in fourteen daily newspapers. The sources identified 61% of the stories with errors. Subjective errors were considered more serious than objective, and as errors increased, credibility with the sources decreased. The study found that circulation had no significant correlation with errors or credibility.

In a similar study of television accuracy, Hanson and Wearden (2004) sent video stories from five Cleveland television stations to 100 sources who had at least five-second sound bites in the stories. Of the 83 respondents, 41% said the sound bite was an incomplete portion of the interview. The sources were asked to complete a semantic differential scale about the quality of the stations. These were combined into a scale, with a resulting .82 correlation ($p < .01$) with the lack of errors.

Only a few of the content analyses have addressed the idea of quality online, although they did not use the term. Carpenter (2008) studied source diversity by content analyzing articles on seventy-two online citizen journalism sites and compared them to articles on fifty daily newspaper sites. Random sampling resulted in 480 citizen journalism articles and 482 newspaper online articles. She studied the number of sources, type of sources (official and unofficial), gender of source, and whether stories about controversies had multiple viewpoints. The newspaper stories had more than twice the number of sources per story (3.64 vs. 1.37). Newspaper stories were more likely than citizen journalism stories to feature official courses, and newspaper stories were more likely to feature both male and female sources than citizen stories. Citizen stories were more likely to present one viewpoint than did the newspaper stories. Generally, online newspaper stories had more source diversity than citizen journalism stories.

Two other studies (Fico, et al., 2013b; Lacy et al., 2010) found that citizen journalism sites varied too much from newspapers to be substitutes, but they could serve the community as supplements to traditional news outlets. One of the studies (Fico, et al., 2013) examined the coverage of local government and found that citizen journalism sites provided less coverage of city councils and county commissioners than did daily and weekly newspapers.

John McManus created an effort in 2000 at Grade the News (www.gradethenews.org) with the goal of using content analysis to measure the quality of local media in the San Francisco Bay Area. In addition to studies conducted by McManus and staff, the site had instructions for visitors to apply a simple protocol to their local news outlets (http://www.gradethenews.org/feat/scoring.htm). The protocol, named the Scorecard, has five variables for assessing quality (topic, knowledge impact, number of named sources, gender diversity of named sources, and, for TV, apparent race/ethnicity diversity). In addition to the content analysis studies, the website carried commentary and qualitative evaluation.

The Grade the News site has not been updated in recent times because of lack of support. Over the years, it was supported by San Jose State University, Stanford University, KTEH Public Television Station, and the Wallace Alexander Gerbode Foundation. In addition to the Grade the News website, McManus has authored three books (McManus, 1994, 2009, 2012a, 2012b) related to news quality and media literacy.

Content Analysis Limitations

Content analysis has been useful in exploring journalism quality because it is unobtrusive, and it can examine content across time. As with all methods, it has limitations. First, the content analysis protocol must reflect standards established in other ways, which in turn require a perspective (news consumer, journalist, academic, etc.). This requirement does not seriously reduce the use of content analysis because a complex protocol could include more than one perspective. However, human content analysis requires coders spending an extensive amount of time evaluating representative samples of journalism content. The use of computers, while less expensive and with greater coder reliability, has limitations as well. Inexpensive software systems that can provide a valid classification based on context and meaning are in their infancy. Even with these software systems, there is an issue of the selecting the standards that lead to the algorithms, and there is the issue of the validity and reliability of programmers. Reliability of software would require that two or more programmers set up the software to yield the same results.

In addition to the number of human hours coding, cost of content analysis has traditionally reflected the expense of generating a representative sample. The availability of journalism content online has reduced the cost of acquiring physical journalism products (e.g., paper, CDs and DVDs). If, however, human coders are used, the issues of how much content is needed during a particular time period must still be resolved. The traditional approach to solving the problem is to use the content creation patterns of media organizations to develop stratified sampling processes that would reduce the number of content units in a sample while maintaining representation. For examples, see Stempel, 1952; Riffe, et. al., 1993, Riffe et. al., 1996; Lacy, Riffe and Randle, 1998.

While these sampling techniques have been used effectively, few studies have examined stratified sampling of online content. For example, one study (Wang and Riffe, 2010) examined sampling of *The News York Times* online, but the *Times* is not representative of the newspaper population. There is no reason to believe the same sampling rules would apply to local news sites. Sampling studies similar to those conducted by Riffe, Lacy, and colleagues should be conducted to examine if efficient sampling strategies can be applied to web and mobile content.

Audience Preferences

Audience preference has been measured through surveys and by examining the relationship between non-audience measures of quality and consumer metrics (circulation, penetration, ratings, etc.). More recently, of course, audience preference has begun to be measured through digital analytics of online behavior, from page views to time per story to sharing.

Surveys

Surveys rely on the ability of people to identify and explain the reason they consume news, and the correlation between quality measures and consumer metrics rest on the assumption that people are purposeful in their news consumption and they match their preferences with content (Lacy, 1989). Surveys of audience members about quality have been varied in sample and in content definitions. Hawley (1992) studied thirty churn household (households that would take and drop newspaper subscriptions) during a two year period and found that the primary reason for dropping was dissatisfaction with content, but the particular content that caused them to drop varied with the household.

During the late 1970s and early 1980s, the American Newspaper Publishers' Association commissioned dozens of research studies by scholars through the ANPA News Research Center (ANPA News Research Report 1-40, 1977-1983). Although the term "quality" was not used, many of the projects aimed to identify the types of content newspaper readers found valuable. The results are too numerous to summarize here, and they may no longer be applicable to today's news environment. However, they used multiple methods (survey, content analysis, in-depth interviews, and circulation figures) to identify types of content important to readers.

Meyer (2004) published a model that states that increases in content quality lead to increases in social influence and credibility, which in turn lead to increases in circulation. To test a portion of the model, he used the results of an annual Knight Ridder reader survey about the credibility of twenty-one newspapers. He found a positive correlation between credibility and the ability of newspapers to maintain their penetration levels.

McCombs, Mauro, and Son (1988) replicated an earlier study (McCombs and Mauro, 1977) by interviewing 357 readers and using a content analysis to identify variables that influence the attention news item received. Item attention was measured as readers were asked if they had noticed, read some, or read most of specific items. Although the strength of the relationship varied with the level of item attention, the best predictors of attention were location of the item on a section front, being written by local staff, being about sports, and having a photograph with the item. The second study was consistent with the first.

Using the typology suggested by Denton (1999), Thorson, et al. (2008) took a survey of about 400 readers in each of a newspaper group’s twenty-seven markets for about 12,000 responses. Quality was measured through seventeen Likert-like scales that addressed types of content. The scales were reduced to five factors using exploratory factor analysis. They reported ten-percent of print readership variance was explained by the five factors with two factors (local news and helps me in my daily life) accounting for most of that variance. In general, quality indicators accounted for little variance in website use, but at one newspaper where the company invested in improving website quality, the five quality factors, particularly “helps me in my daily life,” accounted for seven-percent of the variance in website use.

In a study that surveyed both readers and editors about quality journalism, Gladney (1996) found both agreement and disagreement about what constituted organizational quality and content quality. He received responses from 251 editors at large, medium, and small dailies, as well as from weeklies. He was able to compare these to 291 reader responses from the same markets as the newspapers. The results showed more differences when the rating scales were used than when the rankings of characteristics were used. Looking at the rankings for *organizational characteristics*, the readers agree with the editor on seven of the nine characteristics (with rankings) —1. integrity, 2. impartiality, 3. editorial independence, 5. editorial courage, 6. community leadership, 7. staff professionalism, 9. influence. Disagreement was on staff enterprise, which editorial ranked 4th and readers ranked 8th, and decency, which editors ranked 4th and readers ranked 8th.

There was more disagreement with *quality content characteristics*, but the Spearman's correlation equaled .60, which is a moderate correlation. The two groups agreed on the top three—1. strong local coverage, 2. accuracy, 3. good writing. There was close agreement with three other characteristics—news interpretation (editors-7th and readers-8th), strong editorial page (editors-5th and readers-6th), and community press standard (editors-6th and readers-5th). The greatest disagreement was over lack of sensationalism, which editors ranked 8th and readers ranked 4th. Of course, just what represents sensationalism may differ between the two groups. The last characteristic was comprehensive coverage, which editors ranked 9th and readers ranked 7th.

This is an important study because of its effort to find the areas of agreement and disagreement between the editors and the people they serve. The degree of agreement is encouraging for those who would measure journalism quality because it is the common characteristics were news organizations have the highest probability of serving the people who use their output.

In a case study using choice-based conjoint analysis to measure consumer preference, Kanuri, Thorson, and Mantrala (2014) used a survey of 1,131 newspaper readers to evaluate alternative allocations of news space (preferences). The authors developed recommendation for the newspaper editors about allocation of space to match the respondents' preferences. The editors applied the recommendations and saw an increase in revenue and in positive reader feedback.

Although the term "quality" was not used, the Newspaper Readerships Institute at Northwestern University conducted several studies between 2000 and 2007 that involved large samples of readers. The Newspaper Institute (2001) surveyed 37,000 readers and content analyzed 100 daily newspapers with greater than 10,000 circulation. The study developed what they called the Reader Behavior Score, which combined measures of time spent reading, frequency of reading, and completeness of reading during the week and on Sundays. The results identified nine top areas of content that could be used to build readership (7):

1. **Intensely local, people-focused news.** The specifics of this factor are community announcements, stories about ordinary people and obituaries.
2. **Lifestyle news,** which includes health, fitness and medicine; home, garden and real estate; food; fashion and beauty; and travel.
3. **How we are governed and global relations.** This includes coverage of politics, government, war and international conflict.
4. **Natural disasters and accidents.**
5. **Movies, television and weather.**
6. **Business, economics and personal finance.**
7. **Science, technology and environment.**
8. **Police, crime and the judicial system.**
9. **Sports.** Includes all levels and types of sports.

The study also found that readership could be increased by making the newspaper easier to read and navigate.

The Newspaper Readership Institute (2003) studied 4,400 readers and came up with the idea of the "reader experience." This involves how they feel about the process of consuming information in the newspapers. They identified forty-four distinct experiences that were divided into positive experiences (motivators) and negative experiences (inhibitors). They found that changes in content, service, brand, and marketing could improve the experiences and build readership. The nature of the changes varied with the age group, which means age affects experiences. A couple of the motivators use the terms quality and diversity, but others are consistent with some of the journalism quality characteristics mentioned above. The results are consistent with the news model that argues quality can be defined in terms of user utility (Lacy, 1989).

In a follow-up report, the Newspaper Readership Institute (2004) summarized the finding of the multiple studies to argue that the relationship between newspapers and their readers is based on values, thoughts, and feelings, which they call "experiences." The report argues that newspapers can affect these experiences through content, service, and marketing, but the changes need to be innovative and drastic to attract younger readers.

For years, The Pew Research Center for the People and the Press conducted surveys, first monthly and then weekly, called the News Interest Index, which asked people which stories people followed mostly closely in the news.

One limitation with surveys, of course, is the limitation of human recall. Can people remember accurately what news they consume and the reasons they make those choices? Given this, often the most valuable information from survey research comes from looking at how people's answers might have changed over time, or the longitudinal patterns in the data. If the questions are asked the same way over time, any changes in those responses over the years should be meaningful. Such important trends as declining trust in media or changing consumption habits—which reflect attitudes about quality—are longitudinal studies.

More recently, however, the value of longitudinal data has become challenged by the rapidity of technological change. New technologies require new questions—to probe views and habits about the web, mobile technology social media and more. Not only are longitudinal data patterns possibly interrupted, there can easily become doubts about vocabulary and public understanding of the questions. Many questions that ask people where they get their news, for instance, conflate technology platforms, such as the Internet, with news outlets, such as *The New York Times* or *NBC News*, which today deliver news on many platforms. What constitutes quality for a news gathering organization will raise different issues (size of reporting staff size for instance) than it would for delivery platform might (the nature of a curation algorithm). In addition, consumers may not know rapidly changing terms. The phrase "social media" was not in common use in 2006, although the term "blog" was. Mobile technology did not exist. The word app was not in common use. Nine years later it is unclear what is meant by digital, web or Internet. Perhaps more challenging, long-standing concepts such as "primary news source" are now obsolete. For years, standard questions asked where people got most of their local news or news about international affairs. In the digital age, consumers probably no longer have a primary source for most local or even national news but consume news from many sources, choosing among different ones because what constitutes quality for one subject (such as traffic) may be different than what constitutes quality or value for another (say analysis of politics).

Rosenstiel, et al. at the Pew Research Center (2011) began to try to probe these new behaviors and the end of the so-called primary news source. They broke up local news into 16 different topics and asked people where they went for each one, and the results cast new light on the question of quality. Local TV news, long considered the dominant source of local news, emerged as the source of choice for just three topics—traffic, weather and breaking news—which reflected the immediacy of live newscasts and multiple broadcast times through the day. The newspaper emerged as the most popular source for eleven of the topics, but those were largely civic news that required beat expertise and larger staffs. Online only sources that incorporated more user generated content emerged as the most popular for news about local clubs, restaurants and cultural events.

More recently, The Media Insight Project, a collaboration of the American Press Institute, the Associated Press, and the University of Chicago's NORC (National Opinion Research Center), has expanded this approach of trying to break news into different topics. In its first collaboration (2014), The Media Insight Project found that topic, indeed, was the most significant factor in determining the outlet people seek for news. Consumers are more likely to go to specialized news sources for subjects like sports, ethnic news and technology, where audiences may have deep and specific interest or knowledge. They look to cable for political news to newspapers (again) for local government as well as cultural and local arts. People did not have a preference for any particular technology—convenience is what mattered and that changed with time of day and context. They also articulated clear ability to distinguish technology (such as mobile phones) from delivery platforms (such as Facebook) from news reporting sources (the *New York Times* or NBC). They also preferred to get news, when possible, from the news organization that gathered it—though increasingly they found that news in social media recommendations.

Other research has tried to combine multiple methods to combat the limitations of any one research method. Costera Meijer (2012) published a meta-analysis of seven of her research projects about journalistic quality in European television news. The methods she used included participant observation, in-depth interview, and content analysis. She concluded that audience members want to value and enjoy journalism. She sees a change in the relationship between audience, journalism, and the journalists. She identified three democratic tasks of journalism (4):

1. *Participation*: News users value journalism more when it makes better use of their individual and collective wisdom.
2. *Representation*: News users value journalism more when it takes their concerns and views into account.
3. *Presentation*: Valuable journalism distinguishes itself by a captivating presentation, a gratifying narrative and layered information.

Indirect Measures of Audience Preferences

Several studies have examined the connection between quality measures and audience metrics. A few of these are mentioned above (Lacy & Fico, 1991; Lacy & Sohn, 1989; St. Cyr, et. al., 2005). In addition, Lacy and Martin (1998) compared Thomson Newspapers, as producers of low quality daily newspapers, with a matched random sample of other dailies and used statistical controls. They concluded that Thomson newspapers saw a sharp decline in circulation and penetration between 1980 and 1990 compared to the control group because they were members of the Thomson group.

Stone, Stone, and Trotter (1981) asked a panel of experts to evaluate quality of newspapers and concluded that as quality increased, circulation increased. Cho, Thorson, and Lacy (2004) used a list of dailies with increasing quality identified in *Editor & Publisher* magazine as a measure of quality and then created a randomly sampled control group of similar dailies from around the country. They found that the quality newspapers had higher circulation increases than did the random set of control dailies.

These approaches are almost certain to grow as the data available begins to grow. The American Press Institute is just beginning to assess the data in its Metrics for News program from some 40 publications in which it has full data sets that combine all their metrics and API's journalism meta tags. *The New York Times* has just created an audience development unit that will use data in deeper ways. New publications such as Buzzfeed, Medium, Upworthy, Vice, and Vox are in many ways as much data companies as content publishers, with data scientists in senior positions.

Limitations of Audience Preference Approach

As with content analysis, the audience preference approach has its advantages and limitations. Surveys have the advantage of tapping into the reasons news consumers access journalism and can sometimes be the best way to obtain the consumers' perspective. However, this requires samples that are representative of the groups being studied, which is usually expensive. In addition, closed-ended questions, which standardize responses, must anticipate the attributes that define quality. Just how those attributes are selected can affect the validity of the survey results. Another possibility is to allow open-ended responses and content analyze those results. (The Media Insight Project uses this approach as much as possible). But this also could be very expensive for a representative sample.

The use of consumer metrics such as circulation, time per visit to a website, social sharing and more, to represent audience preference and correlating with quality has appeal, but as reported above (Lacy & Fico, 1991), quality explains a limited amount of variance in such metrics. Variance among respondents can create difficulty in settling on a set of quality attributes whether using surveys or consumer metrics. Lacy and Simon (1993, 38), using the content analysis from Lacy(1986) reported that quality from a professional perspective was positively correlated with circulation but negatively correlated with penetration. This suggests that penetration might be related to breath of content rather than investment in a more narrow set of attributes that would be considered quality from the journalists' perspective.

On top of these problems, people may not consume information the same way on different platforms. Some evidence, for instance, suggests that people "read" differently in print than online. Publishers comparing their print-replica reading patterns data versus their regular web data have found that readers look at a wider variety of stories in print, though they may actually read more deeply on versions designed for the web. Mobile interaction, with smaller screens and technology that is more suited to one action at a time, may be different still.

Expert Judgments of Quality

The use of experts (professional journalists or scholars) to determine quality has been used both with and without other methods. Bogart (1989) surveyed editors and came up with a list of quality attributes. Meyer and Kim (2003) replicated that survey a decade later and found a high correlation in the responses. Burgoon, Burgoon, and Atkin (1982) used the Bogart results as the foundation for surveying journalists, who ranked the top quality elements as follows: accuracy, in-depth reporting, impartiality, investigative reporting, and literary style. Lacy & Fico (1990, 1991) used the Bogart survey to develop a content analysis protocol. Rosenstiel, et al. (2007) used a similar mixing of methods--consulting with news professionals and testing their assessments with consumers both in focus groups and ratings analysis--to see whether consumers could recognize the attributes professionals cited, and whether they tended to see them similarly.

On the other hand, Gladney (1990) surveyed 257 editors at large, medium, and small dailies and at weeklies to compare their responses about organizational standards and content characteristics that represent quality journalism. He found some variation in content elements among the four types of newspapers, but strong local coverage and accuracy ranked either first or second for all types. Good writing ranked either third or fourth at all types. News interpretation averaged fourth for large-daily editors but eighth for all other types of editors. Comprehensive coverage, which meant coverage beyond the immediate market, averaged fifth for the large daily editors and ninth for editors from all other newspaper types. These two differences might represent differences among sizes of communities. Small markets generally have fewer events that would qualify as news in large markets. In other words, the difference may represent variations in news environments rather than professional disagreement.

Other sources of expertise have been used to measure quality. Stone, Stone and Trotter (1981) used a selected group of experts. Cho, Thorson, and Lacy (2003) used the staff decisions from *Editor & Publisher* magazine. White and Andsager (1990) used the number of Pulitzer Prizes won by a newspaper to be a measure of quality in a study of competition. Lacy and Martin (1998) based their decision to treat the Thomson Newspaper group as producing low quality newspapers after the Thomson CEO said as much in a 1993 speech.

API's Metrics for News program involves some similar testing of editor judgments with consumer response. One meta tag used by several of its subject newspapers tags content that editors consider "extraordinary" in quality, and the metrics than reveal whether consumers agree by measuring the reading time, sharing data and number of overall visits to that story. Another tag tests whether what editors believe will go viral or become widely shared is. Certain publishers are also testing what emotional elements in content correlate to sharing (tragic, uplifting, bizarre, informative, controversial). As noted above, the program also tags every story for the level of effort or enterprise that went into it and correlates that to consumer engagement. Enterprise can be seen as a proxy for editor judgment: they put more effort into stories that they both consider more important and that they think consumers will respond to.

Limitations of Expert Judgments of Quality

The idea of defining quality journalism or developing a list of quality attributes by using journalists and scholars has an intuitive appeal. However, it, too, has limits. Journalists and scholars have a perspective that is not shared by all consumers of news, but this difference in perspective can be measured. The same issue of variance mentioned as a limitation for audience members applies here. The Gladney (1990) study showed that not all news managers (and their communities) are homogeneous in their attitudes toward quality attributes. However, the same study suggests some fundamental values are shared across newspaper types.

This process also has the limitations found with all surveys in that it requires a representative sample or a census to accurately summarize the quality attributes from a journalists and/or scholars perspectives. There also is the limitation mentioned from having closed-ended questions, which require a beginning list of potential quality attributes.

Indirect Indicators Approach

This approach toward measuring quality assumes that resources, such as budgets and staff, can be surrogates for quality found in the output. Litman and Bridges (1986) suggested this method and called it the “financial commitment” approach. In their study of how owner concentration affects content, they used four surrogates for quality: number of full-time news/editorial staff, number of news service subscriptions, the absolute weekday lineage for news, and percentage of overall weekday space for news. The study found that lines of news and number of wire services were positively related to a nominal measure of direct daily newspaper competition.

Lacy (1987, 1990) used similar financial commitment measures of performance with a ratio-level index of competition and found a stronger relationship between the quality measures and competition. As a follow-up, Lacy's (1992) model hypothesized that intense competition would lead to financial commitment, which would in turn lead to increased content quality. The increase in content quality would lead to audience utility, and audience utility leads to better market performance (circulation and revenue).

Picard (2000) took a similar indirect approach toward indicators because quality is difficult to measure. He wrote:

> Difficulties in defining quality are especially problematic because the issue of the quality of journalism is not merely a question of increasing the value of a product to consumers. Rather, quality is a central element in achieving the social, political, and cultural goals asserted for journalism in democratic societies (97).

He argued that information gathering and processing activities could be used as a surrogate for measuring quality in content. These activities include: interviews; gathering of information and arranging interviews by telephone; attending events about which stories are written; attending staff meetings, discussions, and training; reading background material; thinking, organizing material; and waiting for information and materials; travelling to and from locations of where information is gathered. This approach is consistent with the financial commitment approach because increasing most of these processes would require increasing funding.

Several studies have examined the relationship between investment and output. Blankenburg and Friend (1994) studied Inland Press Association data for forty-six dailies from 1985 to 1990. They found that an investment in the newsroom was correlated with an increase in circulation but a decrease in profit for the largest circulation dailies in the dataset.

In a study using 1987 Inland Press Association data, Rosenstiel and Mitchell (2004) concluded that four investment variables best predicted revenue growth—newshole as a percentage of overall pages, newshole divided by the number of full-time news/editorial reporters, news/editorial cost per daily average copy, and news/editorial cost per page. They also found that newsroom investment was positively correlated with increased circulation.

In several studies regarding newsroom investment and its impact on newspaper performance, Esther Thorson and colleagues used data available from news organizations. Chen, Thorson, and Lacy (2005) used Inland Press Association data from newspaper with circulation under 85,000 from 1998 to 2002. The number of dailies in each year varied from 258 to 343. They measured financial commitment to the newsroom by using the number of pages in the newshole and the newsroom budget divided by the number of pages in the newshole. The two measured were converted to z-scores and summed. They found positive relationship between the investment index accounted for at least 20% of the variation in adverting revenue per copy and total revenue per copy for all five years.

Taking four of the five years of the dataset used above, Mantrala, et. al. (2007) used econometric modeling to analyze where these newspapers fit on the profit function. They measured news quality investment as the sum of news editorial expenses, newsroom salaries, and miscellaneous expenses. They found an interrelationship between the circulation and advertising market performance. They concluded:

> Second, investments in news quality affect not only subscription sales directly but also advertising revenues through subscriptions indirectly. This result is especially true for smaller-circulation newspapers whose newsrooms and editorial departments tend to be understaffed and overworked. Consequently, our answer to the question "Is good news quality good business?" is a resounding yes. (42)

In a related study, Sridhar, et al. (2011) developed a model of investment for two-sided markets (circulation and advertising) and tested it using data from an unnamed daily newspaper with less than

85,000 circulation. They then replicated the model with another newspaper and concluded that the connection between the two markets must be considered when making decisions for each market. In other words, the advertising market does not subsidize the circulation market.

Tang, et al. (2011) examined 12 years of data from one newspaper to determine the relationship between investment in the newsroom and the performance of the website. They used monthly budget as a measure of investment in the newsroom and found the impact of newsroom investment was indirect. Increases in newsroom investment led to greater readerships and greater revenue. They also found that the investment in newsroom was associated with a greater increase in online advertising revenue than in print advertising revenue.

The studies using indirect measures have been consistent in finding correlations between investment and outcomes (circulation, advertising revenues, profit, etc.). These results are consistent with the financial commitment model (Lacy, 1992), and with Bogart's statement that "a newspaper's investment in its news operation is likely to yield a solid return" (2004, 52). However, this approach sidesteps the actual measuring of content for quality, and it does not seem to be a good way to measure journalism quality defined as service at the community level, which is a utility that accrues beyond the individuals' utility.

Limits of Indirect Measures

Indirect measures assume a connection between the indirect measures (budget or professional activities) and the quality of the product. This assumed linkage has face validity, but as Bogart (2004) points out, the investment is only one of the variables determine the product quality. The research cited above about predicting circulation from quality measures (Lacy & Fico, 1990, 1991; Mantrala, et. al., 2007; Chen, et al, 2005) found only a fraction of variance in outcome variables (circulation, revenue, etc.) was associated with direct quality measures. Other variables contribute to quality as well.

On the other hand, these indirect measures are often easy and inexpensive to access. Further studies that examine the relationship between quality measures and indirect measures using representative samples would establish (or not) the validity of the indirect measures.

Summary of Studies Measuring Quality

Two interesting characteristics of this body of literature are its variance and its consistency. The variance comes from the many ways of conceptualizing and operationalizing the concept of journalism quality and is found in the four basic approaches to that measurement and in the ways of measuring within those approaches. The consistency is found in the relationship between performance measures and the quality measures. Across a variety of journalism quality measures, the studies found a positive connection with increases in audience and other performance measures. This consistency suggests validity of the various quality measures.

In addition, a few of the particular quality measures were used in multiple studies and were consistent across research methods. For example, several studies concluded that strong local coverage was crucial to journalism quality. The mantra of "local, local, local" has been part of the American newspaper industry for almost two centuries. Lately, the term local has transformed into "hyperlocal." The geographic meaning of the term local and hyperlocal, of course, varies with the size and market of the news product. Some newspapers (e.g., community weeklies) cover relatively small communities, such as counties and towns. Metropolitan dailies can cover hundreds or even thousands of square miles. The geographic focus is often mentioned in connection with a story topic. Consumers will say they are interested in local sports, local government, or national politics. Regardless of its variation, geographic and topical emphasis of content are part of quality journalism from both the producer and the consumer perspective.

The presentation of the messages themselves is an element of quality. Both journalists and journalism consumers expect material to be well executed (written, photographed, etc.), accurate and accessible to consumers. Research has explored this area, and measurement is not always easy, especially with subjective accuracy. Other measures also received a wide base of support. For example, in-depth and diverse content is valued by many, and the idea of staff-written correlated with the idea of local coverage.

Of the studies, the ones that incorporate more than one measurement approach seem to be the strongest because support of hypotheses provides support for validity of both types of measures. For instance, the Newspaper Readership Institute studies involved interviewing readers and journalists as well as doing content analysis. The studies did not specifically address the concept of journalism quality, but the elements that are relevant to the concept are consistent with other studies presented here.

Discussion

After all the philosophical discussion and empirical research, what do we know about journalism quality and how to measure it? Does it all boil down to semantic differences in the meaning of quality? How well can we answer Leo Bogart's question about how quality is to be measured? First, we start with the connection between those questions.

The term quality has subjective meaning for various people. We all add connotative meaning to every term we use, but groups of people also share denotative meaning for terms. If we did not, we would be unable to communicate. The trick to defining quality journalism becomes identifying those shared meanings within and among groups. The degree of shared meaning is generally more within groups than across groups. As the research above demonstrates, when you ask journalists to explain the characteristics of quality journalism, they share many of those characteristics. Despite variations, the consistency holds up across different samples and cross time.

Consumers appear to agree less among themselves on the characteristics of quality journalism, which is not surprising. However, the probability is high that identifiable subgroups of consumers would have more shared meanings than a sample of all the consumers in a given market. For example, political junkies are likely to agree more with each other than with someone who only follows sports in the news.

The first step to measurement is the definition of the thing being measured. The second is a typology for classifying variations in the thing defined. This requires a specification of the definition of quality journalism held by a specific group. Once the perspective has been identified and made precise and transparent, debates over what is and is not quality will decline in frequency and intensity. The share quality characteristics within a group can be better determined by interviewing or surveying members of the specific group. This is what Bogart (1989) started with his 1977 survey of editors, which was later replicated (Meyer & Kim, 2003; Burgoon, et al., 1982) and applied to content analysis (Lacy & Fico, 1990).

Of course, if common meaning for quality journalism can be found across groups (journalists and consumers within a given market), as Gladney (1996) found, translating that commonality into measures can be even more powerful in predicting and explaining the relationship between business strategy and quality and between quality and individual and community problem solving.

Just as the quality characteristics discussed in the defining journalism section tended to converge, so did the operational measure of quality used in research projects. However, most of the measures used in research projects tended to emphasize the individual consumer level and represent bundles of information from news outlets. Few of the studies examined quality from the community level. A few studies used source diversity but tended not to apply that beyond the outlet level.

Despite the consistency of journalism quality characteristics and consistent use of some measures from the individual perspective, there are a number of questions that need addressing.

1. What measures can be developed for determining the quality of journalism's contribution to the community? How can we measure quality of source diversity, content diversity, and exposure

diversity? What other measures of community-level journalism quality can be developed? Can these be twinned with measures of whether more robust journalism, or even individual efforts at robust journalism—such as a certain number of exposes conducted each year—correlate to better, cleaner, or more effective civic institutions or economic health?

2. Can the measures developed for newspapers and television be applied to news delivered through mobile and the Internet? A few content analysis studies have applied the same measures to web sites that have been used for traditional media, but research needs to be conducted about consumer and professional perspectives toward differences between traditional and digital news media.
3. What is the role of visual presentation or diverse story telling styles on evaluations of journalism quality? Some scholars have mentioned the role of presentation, but it has not been explored extensively. The API Metrics for News program finds that stories that contain photos average nearly double the engagement online as those that do not. There is evidence that multi-media presentations are more successful than traditional narratives. But there are not enough of them measured in the program to draw firm conclusions. These questions are particularly important because of the increasing use of video and multimedia online and with video. How does the use of visual elements online and with mobile affect professional and consumer perceptions of journalism quality? And how do these different styles of presentation impact audience engagement? Given that any publisher, when operating in a digital environment, is no longer a captive of the limits of one medium, this question is more vital than it has ever been.
4. These studies indicate that there is overlap between perspectives of consumers and journalists. The overlap provides the best measures of quality, but how to deal with those measures that do not overlap remains unresolved. If consumers disagree with journalists about characteristics of quality, how should journalists adjust to the difference? The good news is that the ability to match content analysis with analytics allows a stronger way to know this than ever before.

5. To what end will we use the measures of journalism quality? The goal may affect the appropriate measures to use. Are some measures better for examining the impact of quality on performance? Are other measures better at examining the impact of antecedents (business models, competition, etc.) on quality output?
6. To what degree do consumer preferences and perceptions of quality vary with age, income, gender and other background variables?
7. News is now consumed on demand, often a story at a time or an issue at a time, rather than in "sittings" at defined times of day. API's research about younger Americans also suggests that the ability for consumers to dive deeper into a subject, to access links, see other pieces on the subject, is an important part of utility and value for consumers. The concept that a subject is covered with a main story and perhaps a sidebar each day, with the background material embedded in the body of the story, is an artifact of print, broadcast and radio. Online those elements can be atomized and made more accessible. Links become footnotes. Do these also drive engagement and become in effect elements of Transparency and thus quality?
8. On pages 28 and 29 we referred to seven elements of quality that appear often in academic's and professionals' discussion of journalism quality. Are there important quality elements missing from this list? Should some of these elements be dropped?

Recommendations

This paper is designed to inspire more discussion, not provide definitive answers, and the questions above are ones that, given the rapid and intense changes in the media landscape, deserve more thought. The discussion will be more focused, however, if it includes some initial recommendations and a review of the literature and different approaches provides the chance to identify consistent and enduring ideas about quality going back decades that seem to resonant while moving us ahead. Based on this recognition, we throw out the following:

Diversity, Range and Depth: Nearly every effort to think about quality has involved some concept of diversity, although it has had different names, such as equity and openness. These concepts encompass the diversity that exists inside stories, within publication offerings, and among outlets in a community. In these ways, it offers a strong place to start examining quality.

By measuring the number or range of ***news sources cited, viewpoints*** and ***story topics***, we gain a sense of the richness of an outlet's reporting, its intellectual independence, and a glimpse of whether there is an overriding bias, and a sense of the effort put into the work. If there is a way to measure enterprise, or types of story, this would provide another means of assessing whether the news organization is doing more or less to contextualize stories (Commission on Freedom of the Press) that is to add focus (PEJ) or orientation (as it is referred to by some scholars).

We would add two newer elements to these metrics for Diversity, Range and Depth. By adding ***diversity of storytelling styles***, including narrative, non-narrative and multi-media, we can get a proxy to determine if a publication is making use of the potential of new technology and trying to be creative about the best way to make use of new tools such as data visualization. This would include the presence of ***links and other elements of background, and the ability to learn more.***

By ***adding diversity of story length***, we add another a more subtle but important dimension; whether a publication is covering issues at different levels becomes another kind of diversity of its own, one that would broaden the publication's appeal.

We can get a sense of the richness of journalism available to citizens by community by examining the ***diversity of outlets*** at the community level and then examining the same metrics for each outlet.

Transparency: How candid and forthcoming is an organization about where and how they got their information? This is involves the degree of effort made to share evidence with audiences, to be specific about sourcing, and to reveal the efforts made to verify the news. In addition, transparency involves how candid a news organization is about what it does not know or cannot verify. Transparency gets at the honesty and accuracy that various theorists discussed and also touches on context or orientation (Spurk, Lopata, & Keel, 2010).

Interactivity: A need for a metric directed at the mechanisms that citizens use to interact with the news is suggested by the concepts of news as open, a forum for citizens, and as a way to connect with community. We would recommend metrics that capture the level of what we might call rich and civil interactivity. This doesn't mean simply having a comments section, but having various methods and making sure that they have an element of being moderated so that the conversation is civil constructive and inclusive.

Freedom, Accountability and Independence: The various discussions that the press should be independent, free and would hold those in power accountable is fundamental in American concepts of a free press, but are embedded in other cultures as well. Having ways to measure ***the level of watchdog accountability reporting***, and again ***range of viewpoints***, would get at those fundamentals.

These metrics, which examine content, are all drawn from theories developed by different distinct groups. Add the examination of these metrics to audience engagement data taken from different metrics, from digital analytics to viewership, readership and listenership, would add the third element of audience response. Combine them into an analysis and we would have a robust sense of quality drawn from all four dimensions—theoretical constructs, product and community level measures, and audience engagement as proof of their resonance.

References

Adams, W.C. (1980). Local television news coverage and the central city. *Journal of Broadcasting*, *24*(1), 253-265.

American Newspaper Publishers Association (1977-1983). ANPA News Research Reports, No. 1-40. Washington, DC.

Anderson, P.J. (2014) Defining and measuring quality news journalism. In Anderson, P.J., Ogola, G., & Williams, M. Eds. *The future of quality news journalism*. New York: Routledge, 7-34.

Bernstein, J.M., Lacy, S., Cassara, C., & Lau, T.Y. (1990). Geographic coverage by local television news. *Journalism Quarterly, 67*(4), 663-71.

Bernstein, J. M., & Lacy, S. (1992). Contextual coverage of government by local television news. *Journalism Quarterly*, *69*(2), 329-340.

Blankenburg, W. B., & Friend, R. L. (1994). Effects of cost and revenue strategies on newspaper circulation." *Journal of Media Economics 7*(2), 1-13.

Bogart, L. (1989). *Press and the public: Who reads what, when, where and why in American newspapers*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bogart, L. (2004). Reflections on content quality in newspapers. *Newspaper Research Journal*, 25(1), 40-53.

Burgoon, J. K., Burgoon, M., & Atkin, C. K. (1982). *The world of the working journalist*. Newspaper Readership Project. New York: Newspaper Advertising Bureau.

Carpenter, S. (2008, April 5). *U.S. online citizen journalism and online newspaper stories: A content analysis of source diversity*. Paper presented at the International Symposium on Online Journalism annual conference, Austin, TX.

Charnley, ? (1936). Preliminary notes of a study of newspaper accuracy. *Journalism Quarterly*, *13*(4), 394-401.

Chen, R., Thorson, E., & Lacy, S. (2005). The impact of newsroom investment on newspaper revenue and profit. *Journalism & Mass Communication Quarterly*, *82*(3), 516-532.

Cho, S., Thorson, E., & Lacy, S. (2004). Increased circulation follows investment in newsroom. *Newspaper Research Journal*, *23*(4), 26-39.

Commission on Freedom of the Press. (1947). *A free and responsible press: A general report on mass communication: Newspapers, radio, motion pictures, magazines, and books*. R. D. Leigh, ed. Chicago: University of Chicago Press.

Costera Meijer, I. (2012). Valuable journalism: The search for quality from the vantage point of the user. *Journalism*, *14*(6), 754-770.

del la Piscina, T. R., Gorosarri, M. G., Aiestaran, A., Zabalondo, B., & Agirre, A. (2014). Quality journalism in times of crisis: An analysis of the evolution of European reference press (2001-2012). *Revista Latina de Communicación Social*, *69*, 248-274.

Denton, F. (1999). Chapter 3: The dimensions of local news*. The local news handbook*. Columbia, MO: American Society of Newspaper Editors http://www.asne.org/index.cfm?ID=2908

Diakopoulos, N., & Essa, I. (2008). An annotation model for making sense of information quality in online video. *3rd Conference on the Pragmatic Web*, September 28-30.

Fico, F., & Freedman, E. (2008). Biasing influences on balance in election news coverage: An assessment of newspaper coverage of the 2006 US Senate elections. *Journalism & Mass Communication Quarterly*, *85*(3), 499-514.

Fico, F., Freedman, E., & Love, B. (2006). Partisan and structural balance in newspaper coverage of US senate races in 2004 with female nominees. *Journalism & Mass Communication Quarterly*, *83*(1), 43-57.

Fico, F., Lacy, S., Baldwin, T., Bergan, D., Wildman, S.S., & Zube, P. (2013a). Newspapers devote far less coverage to county government than to city government. *Newspaper Research Journal*, *34*(1), 104-111.

Fico, F., Lacy, S., Wildman, S.S., Baldwin, T., Bergan, D., & Zube, P. (2013b). Citizen journalism sites as information substitutes and complements for newspaper coverage of local governments. *Digital Journalism*, *1*(1), 152-168.

Fico, F., Lacy, S., Baldwin, T., Bergan, D., Wildman, S.S., & Zube, P. (2012). Commentary about government both broad, diverse. *Newspaper Research Journal*, *33*(2), 100-114.

Gertler, M. (2013). Meaning-generating propositions of reality by media: Quality attributes and functions of journalism. *Journal of Information, Communication and Ethics in Society*, *11*(1), 2013.

Gibbons, M. (DATE). Newspaper quality, content and competition in New Zealand. *Pacific Journalism Review*, *20*(1), 181-191.

Gladney, G. A. (1990). Newspaper excellence: How editors of small and large papers judge quality. *Newspaper Research Journal*, *11*(2), 58-72.

Gladney, G. A. (1996). How editors and readers rank and rate the importance of eighteen traditional standards of newspaper excellence. *Journalism & Mass Communication Quarterly*, *73*(2), 319-331.

Hanson, G., & Wearden, S.T. (2004). Measuring newscast accuracy: Appling a newspaper model to television. *Journalism & Mass Communication Quarterly*, *81*(3), 546-558.

Hawley, M. (1992). *Dropping the paper: Losing newspaper loyalists at the local level*. Athens, GA: James M. Cox Institute for Newspaper Management Studies.

Kanuri, V.K., Thorson, E., & Mantrala, M.K. (2014). Using reader preferences to optimize news content: A method and a case study. *The International Journal of Media Management*, *16*(2), 55-75.

Kovach, B., & Rosenstiel, T., *The elements of journalism: What journalists should know and the public should expect*, (2001, 2007, 2014). City?: Three Rivers Press

Lacy, S. (1986). The effects of ownership and competition on daily newspaper content. Ph.D. dissertation, University of Texas at Austin.

Lacy, S. (1987). The effects of intracity competition on daily newspaper content. *Journalism Quarterly*, *64*(2&3), 281-290.

Lacy, S. (1989). A model of demand for news: Impact of competition on newspaper content. *Journalism Quarterly*, *66*(1), 40-48, 128.

Lacy, S. (1990). Newspaper competition and number of press services carried: A replication. *Journalism Quarterly*, 67(1), 79-82.

Lacy, S. (1992). The Financial commitment model of news media competition. *Journal of Media Economics*, *5*(2), 5-21.

Lacy, S. (2000). Commitment of financial resources as a measure of quality, in Robert G. Picard, ed., *Measuring media content, quality and diversity: Approaches and issues in content research*. Turku, Finland: The Media Group, Business and Research Development Centre, Turku School of Economics and Business Administration, 25-50.

Lacy, S., Duffy, M., Riffe, D., Thorson, E., & Fleming, K. (2010). The substitutability between citizen journalism and daily newspaper sites. *Newspaper Research Journal*, *31*(2), 34-46.

Lacy, S., Fico, F.G., Baldwin, T., Bergan, D., Wildman, S.S., & Zube, P. (2012). Dailies do the 'heavy lifting' in government news, despite cuts**.** *Newspaper Research Journal*, *33*(2), 23-39.

Lacy, S., & Martin, M. (1998). High profits and declining circulation: A study of Thomson Newspapers during the 1980s. *Newspaper Research Journal*, *19*(3), 63-76.

Lacy, S., Riffe, D., & Randle, Q. (1998). Sample size in multi-year content analyses of monthly consumer magazines. *Journalism & Mass Communication Quarterly*, *75*(2), 408-417.

Lacy, S., & Simon, T. F. (1993). *The economics and regulation of United States newspapers*. Norwood, NJ: Ablex Publishing.

Lacy, S., & Sohn, A.B. (1990). Correlations of newspaper content with circulation in the suburbs: A comparative case study. *Journalism Quarterly,* 67(4), 785-93.

Lacy, S., Wildman, S.S., Fico, F., Bergan, D., Baldwin, T., & Paul Zube, P. (2013). How radio news uses sources to cover local government news and factors affecting source use. *Journalism and Mass Communication Quarterly*, 90(3), 457-477.

Litman, B. R., & Bridges, J. (1986). An economic analysis of daily newspaper performance. *Newspaper Research Journal*, *7*(3), 7-26.

MacBride, K., & Rosenstiel, T. eds. (2013). *The new ethics of journalism: principles for the 21st century*. Los Angeles: Sage.

Maier, S.R. (2005). Accuracy matters: A cross-market assessment of newspaper error and credibility. *Journalism & Mass Communication Quarterly*, *82*(3), 533-551.

Mantrala, M. K., Naik, P. A., Sridhar, S., & Thorson, E. (2007). Uphill or downhill? Locating the firm on a profit function. *Journal of Marketing, 71*(April), 26–44.

McCombs, M.E., & Mauro, J.B. (1977). Predicting newspaper readership from content characteristics*.* *Journalism Quarterly*, *54*(3), 3-7.

McCombs, M.E., Mauro, J.B., & Son, J. (1988). Predicting newspaper readership from content characteristics: A replication. *Newspaper Research Journal*, *10*(1), 25-30.

McManus, J. (1994). *Market-driven journalism: Let the citizen beware?* Thousand Oaks, CA: Sage.

McManus, J. (2009). *Detecting bull: How to identify bias and junk journalism in print, broadcast and on the wild web*. The Unvarnished Press.

McManus, J. (2012a). *Don't be fooled: A citizen's guide to news and information in the digital age*. The Unvarnished Press.

McManus, J. (2012b). *Detecting bull: How to identify bias and junk journalism in print, broadcast and on the wild web*, 2nd ed. The Unvarnished Press.

McQuail, D. (1992). *Media performance: Mass communication and the public interest*. Sage.

McQuail, D. (2005). *McQuail's mass communication theory*, 5th ed. London: Sage.

Media Insight Project (2014). *Understanding the personal news cycle*. http://www.americanpressinstitute.org/publications/reports/survey-research/personal-news-cycle/

Merrill, J. C. (1968). *The elite press: Great newspapers of the world*. New York: Pitman.

Merrill, J. C., & Lowenstein, R. L. (1971). *Media, messages, and men: New perspectives in communication*. D. McKay Company.

Merrill, J. C., & Odell, S. J. (1983). *Philosophy and journalism*. New York: Longman.

Meyer, P. (2004). The influence model and newspaper business. *Newspaper Research Journal*, *25*(1), 66-83.

Meyer, P., & Arant, M. D. (1992). Use of an electronic database to evaluate newspaper editorial quality. *Journalism Quarterly*, *69*(2), 447-454.

Meyer, P., & Kim, K. H. (2003). *Quantifying newspaper quality: I know it when I see it*. Paper presented at the Association for Education in Journalism and Mass Communication convention, Kansas City, KS, July.

Napoli, P.M. (1999) Deconstructing the diversity principle. *Journal of Communication, 49*(4), 7-34.

Napoli, P.M. (2003). *Audience economics*: *Media institution and the audience marketplace*. New York: Columbia University Press.

Newspaper Readership Institute (2001). The power to grow readership. Northwestern University.

Newspaper Readership Institute (2003). The newspaper experience study. Northwestern University.

Newspaper Readership Institute (2004). Reaching new readers: Revolution, not evolution. Northwestern University.

Picard, R. G. (2000). Measuring quality by journalistic activity, in Robert G. Picard, ed., *Measuring media content, quality and diversity: Approaches and issues in content research*. Turku, Finland: The Media Group, Business and Research Development Centre, Turku School of Economics and Business Administration, 25-50.

Purcell, K., Rainie, L., Mitchell, A., Rosenstiel, T., and Olmstead, K., (March 1, 2010), Understanding the participatory news consumer. Pew Research Center.

Riffe, D., Aust, C. F., & Lacy, S. (1993). The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly*, *70*(1), 133-139.

Riffe, D., Lacy, S., & Fico, F. G. (2014). *Analyzing media messages: Using quantitative content analysis in research*, 3rd ed. New York, Routledge.

Riffe, D., Lacy, S., Nagovan, J., & Burkum, L. (1996). The effectiveness of simple and stratified random sampling in broadcast news content analysis. *Journalism Quarterly & Mass Communication Quarterly*, *73*(1), 159-168.

Rosenstiel, T., Just, M., Belt, T., Pertilla, A., Dean, W., Chinni, D., (2007) *We interrupt this newscast: How to improve local news and win ratings, too.* Cambridge University Press, 8-30.

Rosenstiel, T., & Mitchell, A. (2004). The impact of investing in newsroom resources. *Newspaper Research Journal*, *25*(1), 84-97.

Rosenstiel, T., Mitchell, A., Purcell, K., Rainie, L. (2011) How people learn about their local community, Pew Research Center.

Rosenstiel, T., Sonderman, J, Thompson, T. Benz, J., Sonderman, J, Swanson, E. (2015) How Millennials Get News: Inside the Habits of America's First Digital Generation*,* Media Insight Project

Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass Communication & Society* *3*(1), 3-37.

St. Cyr, C., Lacy, S., & Guzman, S. (2005). Circulation increases follow investment in newsrooms. *Newspaper Research Journal*, *26*(4), 50-60.

Schönbach, K. (2000). Factors of newspaper success: Does quality count? A study of German newspapers. in Robert G. Picard, ed., *Measuring media content, quality and diversity: Approaches and issues in content research*. Turku, Finland: The Media Group, Business and Research Development Centre, Turku School of Economics and Business Administration, 85-95.

Schönbach, K., ed.. 1997. *Zeitungen in den neunzigern: Faktoren ihres erfolgs. 350 Tageszeitungen auf dem pruefstand.* Bonn: ZV.

Simon, T., Fico, F., & Lacy, S. (1989). Covering conflict and controversy: Measuring balance, fairness and defamation in local news. *Journalism Quarterly, 66*(2), 427-434.

Spurk, C., Lopata, S., & Keel, G. (2010). Measuring the 'democratic' quality of radio news: Experiences from Liberia. *Journal of African Media Studies*, *2*(2), 173-192.

Sridhar, S., Mantrala, M. K., Naik, P. A., & Thorson, E. (2011). Dynamic marketing budgeting for platform firms: Theory, evidence, and application. *Journal of Marketing Research*, 48 (December), 929–94.

Stempel, G. H., III. (1952). Sample size for classifying subject matter in dailies. *Journalism Quarterly*, *29*(2), 333-334.

Stone, G. C., Stone, D. B., & Trotter, E. P. (1981). Newspaper quality's relation to circulation. *Newspaper Research Journal 2*(3), 16-24.

Tang, Y., Sridhar, S., Thorson, E., & Mantrala, M.K. (2011). The bricks that build the clicks: Newsroom investments and newspaper online performance. *International Journal on Media Management*, *13*(2), 107–128.

Thorson, E., Meyer, H., Denton, F., Smith, Jim (2008). *Prediction of print and online newspaper readership from indices of news quality*. Paper presented to the International Communication Association, Montreal.

van Cuilenburg, J. (2000). On measuring media competition and media diversity: Concepts, theories and methods, in Robert G. Picard, ed., *Measuring media content, quality and diversity: Approaches and issues in content research*. Turku, Finland: The Media Group, Business and Research Development Centre, Turku School of Economics and Business Administration, 51-84.

Wang, X., & Riffe, D. (2010, May). An exploration of sample sizes for content analysis of the New York Times web site. *Web Journal of Mass Communication Research*, 20, http://wjmcr.org/vol20.

White, H. A., & Andsager, J. L. (1990). Winning newspaper Pulitzer Prizes: The (possible) advantage of being a competitive paper. *Journalism Quarterly*, *67*(4), 912-919.

Wulfemeyer, K.T. (1982). A content analysis of local television news: Answering the critics. *Journal of Broadcasting*, *26*(4), 481-486.

Zeldes, G. A., Fico, F., Carpenter, S., & Diddi, A. (2008). Partisan balance and bias in network coverage of the 2000 and 2004 presidential elections. *Journal of broadcasting & electronic media*, *52*(4), 563-580.

Zeldes, G.A., Fico, F., & Lacy, S. (2008). Context and sources in broadcast
television coverage of the 2004 Democratic primary. *Mass Communication & Society*, *11*(3), 340-346.