Measuring the Scope of Political Communication

Amber E. Boydstun

Prepared for the Visions in Methodology Conference
Ohio State University, October 2-4, 2008

Abstract

Much of politics is conflict. The form varies—from international bargaining to legislative discourse to presidential campaigns—but one fact at least remains constant: the outcome of a conflict is directly affected by its scope (Schattschneider 1960). In the context of political communication, the scope of a conflict refers to how diffuse or focused a discussion is across its alternative dimensions. For example, the debate on the war on terror—whether played out in media articles, political stump speeches, Congressional attention, or elsewhere—can be wide in scope, discussing the issue in terms of democratization, troop safety, weapons of mass destruction, economic cost, etc. Or it can be narrow in scope, focusing on democratization alone. The scope of communication has been shown to have a significant effect on media attention (Boydstun 2008), policymaking (Bevan 2008, Sheingate 2006, Wolfe 2008), public attitudes (Snow, Soule, and Kriesi 2004), and the balance of power in a policy debate (Gamson and Wolfsfeld 1993). Yet methods of measurement differ widely across these and other analyses, perhaps because little has been written on which measure of scope is most theoretically appropriate. I examine the most likely candidates for measuring the scope of political communication and identify one—Shannon's H information entropy—as the best mathematical representation of how scope operates in theory. Additionally, I discuss how utilizing new advances in automated content analysis—that is, using words as data (e.g., Laver, Benoit, and Garry 2003)—can produce a measure of scope that is not only more efficient than relying on hand-coded data but also more consistent with the latent scope variable we are trying to measure. I demonstrate the use of Shannon's H entropy for measuring the scope of communication using a data set of automated word counts taken from all *New York Times* front-page articles on the U.S. conflicts in Afghanistan and Iraq, 2001–2005. I discuss how the Shannon's H measure produced from these automated word counts best captures our theoretical understanding of how the scope of media coverage on the war has evolved over time. The automated method improves efficiency, reduces the need for subjectivity inherent in the manual approach, increases the amount of information the scope measure can leverage, and thus expands the opportunities to incorporate scope as an explanatory variable (or even a dependent variable) in a variety of research puzzles. For example, this method could be used to test whether the scope of international treaty obligations affects states' compliance; whether the scope of legislative debate affects bill passage; or whether the scope of political campaigns affects vote choice.

Introduction

This paper focuses on a single variable that, while potentially of great consequence in the political system, has received sparse attention in political science. This variable, which I call "scope," represents the degree to which political communication is concentrated on a small number of dimensions of discussion or diversified across multiple dimensions of discussion.

The concept of scope is simple enough: Every piece of political communication (such as a speech, a law, a treaty, an entire legislative agenda, even a photograph or a political cartoon) can be considered in terms of the concentration or diffusion of its substance. At a broad level, we can observe how widely or narrowly the "author" of the piece of communication chooses to set the agenda: Does the Senator giving her reelection speech talk about the environment and nothing but the environment, or does she make her audience dizzy jumping from topic to topic?[1]

At a more fine-grained level, we can observe how wide or narrow the author's discussion is on a specific topic at hand: In talking about the environment, does the Senator focus only on

---

[1] The term "dimensions of discussion" refers to the universe of substantive elements that could (at least hypothetically) receive some amount of attention in a given communication piece. At the broadest level—such as when considering the scope of a political speech—the environment is a dimension of discussion, as is energy, defense, health care, and all other political topics that might receive some amount of attention in the Senator's remarks, on the front page of the paper, on a legislative agenda, or in any other political communication venue of interest. Practically speaking, the total number of dimensions that could be used in a piece of communication depends upon our system of measurement. If we code communication texts by hand according to a pre-defined coding scheme, the total number of possible dimensions must be defined *a priori* within the structure of that coding scheme. But if, for example, we rely on automated content analysis to compile word frequencies by which to calculate a measure of scope, the total number of possible dimensions is set equal to the total number of distinct words (or sets of words with a common root) in the data set.

the issue of endangered species protection, or does she talk also about land conservation, air pollution, deforestation, and so on?[2] And in talking about endangered species, does she define—or frame—the issue only in terms of the moral obligation to protect earth's creatures, or does she discuss other dimensions of the issue as well, such as the financial costs of species protection, the needs of loggers and other laborers in environmentally contested areas to feed their families, and so on?[3] At each of these levels, scope refers to how broadly or narrowly political communication is crafted.

There is good reason to believe that the scope of communication will play a key role in unraveling many political puzzles. In the realm of politics, most items of communication occur against the backdrop of a political body (voters, legislators, world leaders, etc.) deciding what stance to take on a given question. In other words, most political communication occurs in the context of a conflict, be it who to elect for President, what demands to include in an international treaty, or what policy to implement. Thus, whether intended or not, each piece of political communication serves to define, or frame, the conflict at hand by dictating *how many* and *which* aspects of the conflict are considered.

Research on framing has shown how the content of political communication can, under the right conditions, have a significant impact on individual attitudes (Brewer and Gross 2005, Chong and Druckman 2007, Dardis et al. 2008, Nelson et al. 1997, Nelson and Oxley 1999,

[2] Similarly, within the topic of the environment, the issues of deforestation, land conservation, water pollution, and endangered species protection are all dimensions of discussion that could be considered, as are many more we could name.

[3] In discussing the issue of deforestation, each of the perspectives or "frames" that could theoretically be utilized in the debate—such as species protection and the economic needs of logging communities—represents a dimension of discussion.

Wood and Doan 2003), on mass public opinion (Gilliam and Iyengar 2005, Jacoby 2000, Jerit 2008, Kellstedt 2003, Kinder and Nelson 2005, Pollock 1994, Terkildsen and Schnell 1997), on social movements (Gamson and Wolfsfeld 2003, Snow et al. 2004), and on public policy (Baumgartner et al. 2008).

Examining the *scope* of political communication is a natural next step in framing research. As Schattschneider said: "The outcome of all conflict is determined by the *scope* of its contagion" (1960: 2, emphasis in original). Already, studies have demonstrated the importance of scope in a variety of contexts (Bevan 2008, Boydstun 2008, Jones and Baumgartner 2005, Sheingate 2006, Wolfe 2008).

Moreover, scope may prove a fruitful variable for consideration in several other areas of political science. For example, all else being equal, when there is marked variance in the scope of communication used by two Presidential candidates—Candidate A spends her speeches talking about only one or two policy issues, while Candidate B uses her speeches to cover many different issues—does the variance in scope affect public perception of the candidates, or even vote choice? Does the scope of an international treaty—how concentrated it is on a select number of items or spread across many items—have any influence on how likely the treaty is to succeed? Are legislative sessions that handle a restricted set of policy problems more or less productive than legislative sessions that tackle a wide array of issues? Does the scope of news attention to the U.S. conflicts in Afghanistan and Iraq affect how *much* news coverage the war receives? Investigations into questions like these will advance our understanding of important elements in the political process (elections, international treaties, policymaking, media attention, and so on).

Yet any research agenda involving scope depends on having a measure of this complex variable that is both theoretically appropriate and logistically feasible. In this paper, I propose such a measure. I begin with a more detailed discussion of scope, explaining what this under-utilized variable represents and why it matters. Next, I examine the theoretical mechanisms of scope. By unpacking how scope works in the abstract, I develop a rough litmus test by which to assess each measurement option. Next, I examine five methodological candidates for measuring scope, comparing each one to my theoretical understanding of how scope operates. From this discussion, I identify one key measure— Shannon's H information entropy—that I suggest for studies using scope. I also propose utilizing automated content analysis rather than hand coding in the process of collecting the political communication data sets from which scope can be calculated. Finally, I illustrate the use of automated content analysis and Shannon's H entropy to calculate scope in a data set of all *New York Times* front-page articles on the U.S. conflicts in Afghanistan and Iraq, 2001–2005.

What Is Scope and Why Is It Important?

The scope of political communication is the degree to which communication is narrowly focused on a small set of dimensions or widely encompassing of multiple dimensions. We can think of each piece of political communication (or a group of communication pieces occurring in a given venue or time period of interest) as an agenda—a "pie" of attention, if you will. The process of allocating this agenda across one or more policy issues (unemployment, immigration, the war, anti-monopoly regulation, endangered species protection, etc.) is what institutional scholars call "agenda-setting."[4] Looking just at the communication about a specific policy issue (e.g., deforestation), the process of allocating the agenda across one or more perspectives of the issue (e.g., forest protection, logging communities) is called "issue-definition" or "issue-framing."

The allocation of political communication along competing dimensions has long been recognized as an important part of the political process. At both the agenda-setting and the issue-framing levels, the dimensions included in discussion are the ones that get dealt with; these are the issue frames that garner public concern and ultimately receive policy response. Dimensions excluded from the agenda are ignored. In this way, agenda-setting and issue-framing are powerful political mechanisms.[5]

---

[4] In contrast, behavioral scholars use the term agenda-setting to refer to the *effects* of how an agenda is divided across issues on individual attitudes, vote choice, etc. In this context, the term issue-framing is sometimes referred to as "second-level agenda-setting" (e.g., McCombs and Ghanem 2003). This difference in terminology is mostly semantic, as the institutional and behavioral conceptions of these concepts are commensurate.

[5] In the realm of institutional agenda-setting, the scarcity of attention on any communication agenda means that the objective urgency of a policy issue is not a sufficient criterion for attention. An issue must be important *enough*, compared to all other issues, to pass a "threshold of urgency" as defined by the congestion of the agenda (Jones and Baumgartner 2005). Whatever form the agenda takes, it can have a profound influence on public opinion (Behr and Iyengar 1985, Dearing and Rogers 1996, Entman 1989, Iyengar 1991, Iyengar and Kinder 1987, Iyengar and McGrady 2007, Jacoby 2000, Jasperson et al. 1998, McCombs 2004),[5] governmental agendas (Edwards and Wood 1999, Wood and Peake 1998), and public policy (Birkland 1997, Walgrave and Van Aelst 2006). In the realm of issue-framing, research has shown that the public does not respond to real-world problems so much as to the collective definitions society adopts for those problems (Best 1995; Blumer 1948; Hilgartner and Bosk 1988; Spector and Kitsuse 1973, 1977). As Kingdon puts it, "conditions become defined as problems when we come to believe that we should do something about them"

Traditionally, agenda-setting and issue-framing research has focused on the *content* of political communication: which issues or frames are being discussed, and which are being ignored. Yet although scholars have rarely examined the *scope* of political communication, consideration of this single variable offers an important new line of research devoted to understanding how the diversity or uniformity of a political message affects the impact of its content.

When Schattschneider wrote that “the outcome of all conflict is determined by the *scope* of its contagion” (1960: 2, emphasis in original), he was referring to the amount of public attention paid to a debate (i.e., the number of people involved).[6] The more people involved in a debate, the wider the scope of the conflict. Why does the scope of a conflict matter? Because when the scope of a debate expands, the influx of new dimensions (or people) disrupts the status quo line of debate. And with enough expansion, the line of debate can be displaced entirely, such that the previous losing side of the debate becomes the new majority (Riker 1986, Schattschneider 1960).

In the realm of political communication, the scope of a debate refers to how encompassing communication is across dimensions of discussion, not across human beings, but the concept is the same. New dimensions in political communication mean new ways of thinking about an issue and, thus, a greater likelihood of resonating with more segments of the public and

---

(1995). In particular, issue-framing can serve to propel a policy issue above the threshold of urgency needed to obtain a slice of the agenda (Cobb & Elder 1972; Kingdon 1995; Baumgartner & Jones 1993 and 2002).

[6] Explicitly, Schattschneider writes: “The scope of conflict is an aspect of the scale of political organization and the extent of political competition. The size of the constituencies being mobilized, the inclusiveness or exclusiveness of conflicts people expect to develop have a bearing on all theories about how politics is or should be organized” (1960: 20).

displacing the current line of conflict in the public debate. Each time a new dimension is activated in political communication, the line of debate is necessarily redrawn. Sometimes the shift will be minor. Sometimes it will realign the entire debate. But every change in the dimensional structure of a political discussion threatens the status quo. As Schattschneider says, conflict displacement is "*the most devastating kind of political strategy*" (1960, emphasis in original).

In their discussion of policymaking, Baumgartner and Jones talk about how introducing new dimensions (what they call "principles") of an issue debate can alter policy outcomes. "When a general principle of policy action is in place," they write, "policymaking tends to assume an incremental character. When new principles are under consideration, the policymaking process tends to be volatile, and Kingdon's model is most relevant" (1993). In other words, the scope of political communication is a prime vehicle for agenda change. When the scope expands, it softens the ground for a critical juncture of attention redistribution.

In particular, an expanding debate—one that moves from being defined along a single dimension to being defined across a range of dimensions—can lead to a cascade of attention (Baumgartner et al. 2008, Boydstun 2008). For example, an analysis of the *New York Times* front-page agenda suggests that an increase in the scope of media attention to a given policy debate has a significant positive influence on how much media attention the issue receives in the following time period (Boydstun 2008). Consider that in the 1950s discussion about nuclear power was dominated by talk of scientific advancement: the "atoms for peace" frame. But in the late 1960s it was reframed in terms of environmental danger, health risk, and military arms proliferation. This shift in emphasis had an immediate and enormous impact on public perception of and policy response to the issue (Baumgartner and Jones 1993). But what was important about the shift in definition was not merely a swap of one set of frames for another. The new frames realigned the debate not just through their substance but through their sheer number; the scope of the debate had expanded.

The importance of scope has also been examined in the context of policy debates more generally. Snow et al. (2004) find that the scope of a debate strongly shapes public response to that debate. And Gamson and Wolfsfeld (1993) offer compelling evidence that the scope of a debate directly affects the likelihood that the balance of power in the debate will shift. Research in the policy area of capital punishment suggests that the dimensions of a policy debate piggyback, or resonate, off of one another. As the scope of a policy debate expands to include additional dimensions, it makes it easier for even more dimensions to gain traction on the agenda. An expanding debate is an exciting debate, and as the scope of discussion widens, attention increases. Again, the expanding scope of an issue debate is a strong contributing factor in attention cascades (Baumgartner et al. 2008).

Finally, in the realm of legislative attention and policymaking, Sheingate (2006) finds that the scope of a Congressional committee's jurisdiction can, under the right conditions, have a significant effect on how active the committee will be in a given policy issue. Bevan (2008) examines executive and legislative agendas from five different countries and also finds a strong positive relationship between the scope of an agenda and its size. In particular, an expansion in the scope of a political agenda allows newcomers to push their policy concerns to the forefront (Bevan 2008). Wolfe (2008) shows how the scope of the information available in the political environment can have a significant effect on policymaking, both because the concentration or diffusion of the information environment directly influences government responsiveness and because the scope of the information environment can attenuate the signal strength of public opinion and, thus, the impact of public opinion on public policy.

In short, the scope of political communication (in the form of a text, a speech, a legislative agenda, or otherwise) represents a promising and important direction of research spanning several areas of political science. But in order for researchers to utilize scope effectively as a variable of interest, we must first have a good way of measuring it. In the next section, I discuss the theoretical mechanics of scope that together will serve as a litmus test for the different measurement options I will examine.

How Does Scope Work?

In order to select a method for measuring scope, we must first have a more concrete understanding of how scope functions. Because scope is a complex and amorphous latent variable, there is no surefire way of comparing the relative accuracy of different empirical measures. And so, as it should be, we must rely on theory to guide our methods. In this section, I discuss the theoretical mechanisms of scope: how it operates generally, how it should be expected to behave as it changes, and thus what criteria we should consider in selecting a measure.

What do we know about how the scope of political communication operates? Does it behave consistently across different conditions? What would an ideal measure of scope look like? To begin, we can conceptualize scope as extending across a two-dimensional scale.[7] At one

[7] Perhaps future research will extend our understanding of scope to include consideration of a third dimension, such as the signal strength or valence of the piece of communication being studied. But for now two dimensions offer enough to keep us busy.

end, the smallest value of scope represents complete concentration on a single dimension. At the other end, the largest value of scope represents complete diffusion across all possible dimensions. Thus, at a very basic level, we want a measure of scope that increases as scope expands and decreases as scope contracts.

Imagine the positive quadrant of a Cartesian graph. By knowing that we want the lowest value of our measure to represent complete concentration and the highest value to represent complete diffusion, we have tethered the two end points of our theoretical measurement function: one at the intersection of the x and y axes, and one at the farthest upper right-hand point on the graph. How, then, should we connect these two points?

If we believe that scope functions uniformly across the range of its values, then we should search for a linear measure that connects our two points with a simple straight line (slope=1). But scope does not work this way. Consider the speeches over time given by a political candidate. If at some point in time this candidate changes from talking only about a single issue—like endangered species protection—to talking about two issues—say, endangered species protection and Medicare benefits—does this increase in scope have the same weight as if the same candidate at the same pivotal point in time went from talking about *seventy*-one issues to talking about *seventy*-two? Absolutely not. Although in both cases change occurs in the form of a single additional issue, the political significance of moving from an agenda of one to an agenda of two is much greater.

Thus, we need a curvilinear measure of scope. In particular, we need our measure to be concave; that is, more sensitive to increases in scope at the lower levels of its value range and less sensitive at higher levels. In fact, at some point we should expect an accurate measure of scope to plateau, corresponding with the point at which a piece of political communication becomes so saturated with dimensions (i.e., with information) that an increase in scope—even a big one—is barely noticed.

However, between the jump from one to two dimensions and the jump from something like seventy-one to seventy-two, we want the slope of our curvilinear function to be somewhat gradual. Although no other unit increase should produce as great an increase in our measure of scope as the 100% increase from one to two dimensions, depending on the context of our study we probably want the increase from, say, four to five dimensions to carry considerable weight.

Beyond the shape of our ideal curvilinear function, we want our measure to behave well when faced with real data. The measure should be as robust as possible across data sets of varying observation size. An ideal measure would also perform robustly across different studies with varying numbers of possible dimensions at work in the political communication venue being analyzed. Yet our measure must also take the number of possible dimensions into account, such that it differentiates between the importance of increasing from eight to nine dimensions in a theoretical universe of ten dimensions and the importance of increasing from eight to nine dimensions when hundreds of dimensions could be used.

Finally, in the interest of dynamic studies, where multiple pieces of communication are analyzed together to constitute the overall communication signal within each time point, our measure must not be hyper-sensitive to changes in the number of observations over time. For example, a sudden drop in the number of newspaper articles, press releases, or bills about a particular issue should not necessarily influence the calculated scope value. The scope of a policy discussion and the total amount of attention the policy receives are mutually-reinforcing in many ways (Boydstun 2008), but a sudden change in observations can often result from unrelated factors in the system, such as the sudden presence of another hot topic on the agenda.

We want our measure of scope to be attuned to how the potential set of dimensions available in a discussion are used or not used, whether the total size of the discussion occupies ten observations in a given time period or ten thousand.

How Do We Measure Scope?

How then should we measure the scope of political communication? We want a measure that captures the varying importance of scope expansion relative to the current baseline. We want a measure that gives more weight to increasing from one dimension of discussion to two than from seventy-one dimensions to seventy-two but with some gradation in between. We also want a measure that performs consistently across data sets of different observation size and across communication venues of different dimension size. And, for dynamic studies, we want a measure that is resilient to unrelated dips or spikes in the number of observations available in a given time period. In other words, we want a non-linear mathematical function that matches our theoretical understanding of how scope operates.

In this section, I identify and discuss five candidates for measuring the scope of communication: 1) a simple proportion of the dimensions employed relative to the total number of dimensions, 2) Herfindahl- Hirschman Index, 3) Shannon's H information entropy, 3) the eigenvalue of the first factor identified by evolutionary factor analysis, and 4) the amount of variance left unexplained by the first factor identified by the same evolutionary factor analysis. These candidates represent the best options for measuring scope suggested by existing literature.

Importantly, all the measurement options I discuss require an initial data set containing either 1) the results from hand-coding a set of communication pieces using a pre-defined coding scheme, or 2) the results from subjecting a set of communication pieces to automated content analysis. As I will discuss, I propose collecting data by means of the latter method.

Throughout my discussion of these five possible measures, I examine how each behaves in the abstract using simulated data. This abstract demonstration offers a clear look at the important differences between the measures, allowing us to identify one candidate in particular as the best match for our theoretical understanding of scope.

Proportion of Dimensions Used

The most straightforward method for measuring scope is to take a simple count of the number of dimensions used in a piece of political communication (or across multiple pieces of communication within a time period of interest) and then calculate this count as a proportion of the total number of dimensions that could be used. While it will quickly become apparent that this measure is not one we should adopt, it offers a good baseline from which to begin our search. This measure is expressed in the formula below.

Formula 1. Proportion of Dimensions Used

$$P = \frac{\sum_{i=1}^{n} x_i}{N}$$

where:
x_i represents a dimension
N represents the total number of possible dimensions[8]

[8] Again, the total number of possible dimensions depends on our system of measurement. Using hand-coded data, N is defined *a priori* by the coding scheme the researcher develops. Using automated word counts, N equals the total number of distinct words (or sets of words with a common root) that appear in the data set.

Using this measure of scope, if there are 100 possible policy topics the President could discuss in his State of the Union address and he talks about 28 of them, the scope value for that address is set at 0.28. If the news stories in a given month reference 15 distinct frames within a given issue debate that contains 30 theoretical dimensions, the scope value for that issue for that month is set at 0.50.

Figure 1 compares the behavior of all the considered measures across a full range of scope observations in simulated data. In the case of calculating a proportion of the dimensions used, compiling simulated data was very easy. Imagining a hypothetical communication venue with *N* dimensions, I simply constructed an artificial series of observations tracking the proportion of those dimensions (out of 1.0) that received some share of the agenda. Using 100 observations, I set the value of each observation such that it registered one additional dimension than the previous observation, starting at 1 dimension (scope=0.01) and ending at 100 dimensions (scope=1.00).

The behavior of this measure is illustrated as the thin solid line in Figure 1. Unsurprisingly, the result is a perfectly straight line stretching from lower-left to upper-right. Since we have already decided that we need a measure of scope based on a curvilinear function, we can eliminate this measure from our list of candidates.

Herfindahl- Hirschman Index

The first viable candidate to consider is the Herfindahl-Hirschman Index, originally developed by economists to measure the concentration of industries (i.e., monopolies). As shown in the formula below, the Herfindahl Index is calculated by taking the square of the proportion of attention captured by each dimension, summing these squares, and then subtracting the total from one.

Formula 2. Herfindahl-Hirschman Index

$$HHI = 1 - \sum_{i=1}^{n} (p(x_i))^2$$

where:
x_i represents a dimension
$p(x_i)$ is the proportion of total attention the dimension receives

Notice that by subtracting the sum of the squares of the proportions from one, the formula is scaled such that values range from 0 to 1, with zero representing complete concentration (e.g., in terms of industrial economics, that one firm controls an absolute monopoly) and 1 representing complete diffusion (e.g., each industry has an equal share of the market).

While the math is simple, we can see that the Herfindahl Index represents a theoretical shift from the baseline proportion measure considered above. Rather than relying on a raw count or proportion capturing how many dimensions are used in a communication piece, the Herfindahl Index takes into account how the communication was allotted across these dimensions. If twenty dimensions each receive an equal amount of attention in a speech, for example, the scope value calculated here will be very different than if one dimension received the lion's share of attention (perhaps measured in words or minutes) and the other nineteen were only mentioned in passing. Specifically, as the proportion of the agenda given to any single dimension *decreases*—suggesting, at least with regard to that dimension, that attention is being spread more evenly across all dimensions—the Herfindahl Index value *increases*, signaling greater dispersion.

We can observe how the Herfindahl Index behaves as a measure of scope by looking again at Figure 1, this time at the line with white circles at the far left side of the graph. As before, these data points were produced by calculating the Herfindahl Index formula using simulated data. This time, I constructed 100 different sets of observations, each set representing a different allocation of communication across 100 fictitious dimensions. In the first set of observations (represented by the first data point at the lower left-hand corner of Figure 1), one dimension controlled 100% of the agenda ($p_i = 1.00$); in other words, the agenda was shared by only 1% of the dimensions available. In the second set of observations, the agenda was divided equally between two dimensions (or 2% of all dimensions), with each dimension receiving 50% ($p_i = 0.50$), and so on. The final observation seen in the far upper right-hand corner of Figure 1 represents the last set of fictitious data, where all 100 dimensions (100% of those available) share the agenda, with each dimension receiving 1% of attention ($p_i = 0.01$).

Unlike the simple proportion measure we considered above, the Herfindahl Index displays a markedly curvilinear form. However, the sharp increase in values within the first 10% of observations is cause for concern. This measure satisfies our criterion of giving more weight to a change from one to two dimensions, but in fact it gives too much weight. Using this measure, the increase in scope value produced by increasing from nine dimensions to ten is only nominally larger than the increase in scope produced by increasing from 99 dimensions to 100.

Of additional concern is the fact that the Herfindahl Index formula does not control for the number of possible dimensions at work in a given communication venue. There are ways to adjust the formula to account for the number of dimensions,[9] but these measures retain the problem of being over-sensitive to low-level changes in scope and under-sensitive to high-level changes in scope.

---

[9] For example: $HHI^* = \dfrac{1 - \sum_{i=1}^{n} (p(x_i))^2 - \frac{1}{N}}{1 - \frac{1}{N}}$

Shannon's H Information Entropy

Shannon's H Information Entropy formula is a variant of the generic entropy formula, originally developed in the field of thermodynamics to measure the diffusion of heat. Shannon (1948, Shannon and Weaver 1949 and 1971) proposed that human communication can be understood in terms of the concentration and diffusion of the categorical information it contains, and he developed the information variant of the entropy formula in this context. In the field of political science, Shannon's H has arguably been used more often than any other single measure in studies of scope. In particular, this measure has been used to study institutional agenda-setting (Baumgartner et al. 2000, Boydstun 2008), comparative policy attention (Bevan 2008), shifts in agenda volatility (Talbert and Potoski 2002), Congressional committee jurisdiction (Sheingate 2006), and information complexity (Wolfe 2008).

As the equation below shows, Shannon's H is calculated by multiplying the proportion of the agenda each dimension receives by the log of that proportion, then taking the negative sum of those products. Importantly, the log function in this formula uses as its base the total number of possible dimensions in the communication venue under study. Thus, unlike the standard Herfindahl Index, the Shannon's H entropy measure directly accounts for the size of the dimensional universe. In this way, the measure can appropriately differentiate between the use of only nine dimensions in a communication piece where only ten dimensions are available and the use of nine dimensions when hundreds are available.[10]

[10] See footnotes 1 and 8 above.

Formula 3. Shannon's H Information Entropy

$$Entropy = -\sum_{i=1}^{n} p(x_i) * \log_n p(x_i)$$

where:

x_i represents a dimension
$p(x_i)$ is the proportion of total attention the dimension receives
$\log_n p(x_i)$ is the log of the proportion of attention the dimension receives, using the total number of possible dimensions as the base of the log

As with the Herfindahl Index, Shannon's H entropy increases as the proportion of attention that a given dimension receives decreases. In other words, as the spread of attention across all dimensions becomes more equal, or diffuse, entropy increases.

Looking again at Figure 1—this time at the line with black diamonds—we can see that Shannon's H entropy has a much more gradual curvilinear structure than produced by the Herfindahl Index. In order to obtain the values underlying the Shannon's H line, I used precisely the same simulated data series as described above with regard to the Herfindahl Index. The only difference was that, instead of applying the Herfindahl Index formula to the simulated proportions, I applied the Shannon's H entropy formula listed above.

Unlike the Herfindahl Index measure, Shannon's H does an excellent job of giving extra weight to low-level changes in scope while not overly muting the significance of medium- and high-level changes. As Bevan (2008) notes, the Shannon's H entropy measure shows a higher degree of variance and, thus, "is far more sensitive to states of greater scope than the Herfindahl Index."

Evolutionary Factor Analysis: Eigenvalue and Unexplained Variance

Although the Shannon's H entropy measure appears well suited to the theoretical workings of scope, there are two final measures I will consider, both of which are produced by performing evolutionary factor analysis on political communication data over time. As outlined by Baumgartner et al. (2008), the evolutionary factor analysis approach allows researchers to uncover the dimensions being utilized in a set of political communication pieces and to observe how the composition of these dimensions changes over time.

In brief, evolutionary factor analysis is based on the assumption that there are common threads, or "factors," that weave through any related collection of data. When factor analysis is performed on a set of variables it identifies the number of strong factors latent in the data and assigns each variable an "eigenvalue" representing the extent to which that variable contributes to the common factor. Additionally, this method calculates the proportion of the total variance in the data set that is explained (and, thus, not explained) by each of the main factors.

Evolutionary factor analysis utilizes the same analysis but performed dynamically. Instead of looking for common threads within a cross-sectional cluster of data, evolutionary factor analysis looks for threads that weave across a series of consecutive time periods. By performing factor analysis on moving windows of 5 or so time periods—working sequentially forward so that each new window drops the oldest time period from the previous window and adopts the next time period ahead in the series—this method reveals not only which factors are at work at which points in time but also how the composition of each factor changes over time as additional variables begin contributing to the factor and other variables stop contributing (Baumgartner et al. 2008).

Why consider factor analysis as a measure of scope? Because latent semantic research shows that the statistical mechanisms of factor analysis mirror the cognitive processes of language construction and comprehension (Landauer and Dumais 1997). The implication of latent semantic research is that factor analysis is a theoretically appropriate tool for parsing the dimensional structure of human texts. Thus, while Shannon's H entropy appears to be a promising measure of scope, it would be imprudent to accept it as our measure without first considering the measures we can extract from evolutionary factor analysis.

I examine two possible measures of scope in this context: the eigenvalue of the primary factor uncovered and the amount of variance in the data set left *unexplained* by this first factor. Because of the complexities of evolutionary factor analysis, neither of these measures can be summarized as a formula. But we can see how each operates by looking again at Figure 1.

Unlike the first three lines we considered in Figure 1—each of which was generated using "perfect" simulated data (e.g., first one dimension receives 100% attention, then two receive 50%, and so on)—the lines representing the evolutionary factor analysis measures in Figure 1 are based on simulated data using random seeds. Because factor analysis relies on examining variance, the computer code used to perform factor analysis simply will not run on perfect data. Instead, I developed a new simulated data set, this time with each of 100 fictitious dimensions receiving a random amount of attention. I then varied the spread of attention across dimensions in each of 100 observation sets, but the amount of attention that a given dimension received (initially computed as a random number) did *not* change from one observation set to the next. In fact, the only thing I changed in each observation set was to add one additional dimension that had previously received no attention and to set it at its fixed amount of attention. Thus, in the first set of observations, only one dimension received attention. In the second set, two received attention; in the third set, three did; and so on. In the final set, all 100 dimensions received attention (again, with each dimension receiving the exact *amount* of attention it had received in every other observation set in which it was activated).

After developing this simulated data, I subjected it to evolutionary factor analysis using moving 5-month time windows. The eigenvalue and percentage of variance left unexplained for each primary factor are recorded as the two final lines (grey triangles and grey squares) in Figure 1. Surprisingly, both measures perform poorly when compared with the theoretical criteria we outlined in the previous section. The eigenvalue measure turns out to have a functional form almost identical to the baseline measure of the proportion of dimensions used we considered and rejected. In retrospect, this similarity is sensible. Eigenvalues are calculated directly from the variance in the data, and so if changes in variance occur incrementally (as do the changes I instituted in all the simulated data sets shown in Figure 1), the resulting function should be linear. In contrast, the percentage of unexplained variance produces a curvilinear function. However, as with the Herfindahl Index, this measure plateaus so suddenly and at such a low level of scope that it ignores the important variance in scope that occurs between nearly complete concentration and nearly complete diffusion.

Thus, although these two measures based on evolutionary factor analysis held theoretical promise as ways of capturing scope, they do not perform well—even when considered in the abstract context of Figure 1. Having examined all five measures in the abstract, it is clear that one measure alone—Shannon's H information entropy—stands well above the others as a theoretically appropriate measure of the scope of political communication.

Illustration: *New York Times* Coverage of the U.S. Conflicts in Afghanistan and Iraq, 2001–2005

Thus far, my discussion has been restricted to an abstract consideration of how to measure scope. Having identified Shannon's H information entropy as the measure that best aligns with our theoretical understanding of how scope operates, I use this section to illustrate the use of this measure (and, for reference, the other competing measures) in the context of real data. The data set I utilize contains all front-page *New York Times* articles on the U.S. military conflicts in Afghanistan and Iraq from September 12, 2001 through December 31, 2005. The articles in this data set have already been coded by hand according to an intricate coding scheme designed to capture shifting media attention to different dimensions—or frames—of the war debate over time (Boydstun and Glazier 2008). Thus, having already analyzed this hand-coded data elsewhere, I have a solid understanding of how the war debate has unfolded and, thus, what a measure of scope "should" produce when applied to this data set.

I begin with Figure 2, which compares the same measurement candidates shown in Figure 1 (minus the proportion of dimensions used), this time calculated on the data set of *Times* front-page war stories. For each line in this graph, I performed the same calculations described in the sections above. This time, however, instead of using simulated data I used the collection of word counts produced by running the *Times* stories through a simple automated content analysis software program. In a nutshell, automated content analysis is based on the theory that the substance contained in human communication can be captured by performing statistical analysis (ranging from simple word counts to complex algorithms) on the component words of that communication (Laver et al. 2003).

In this case, I collected frequency counts of all the words used in all the *Times* stories about the war. Since I wanted to use months as my units of analysis, I clustered these word counts within each month. For example, if the word "democratization" is used 10 times in one article and five times in another article written in the same month, the total word count for democratization for that month is 15. After excluding non-substantive words from the data set (a, and, the, of, etc.) and clustering common root words (e.g., democracy, democratize, democratization), the resulting dataset represents the core substance of the communication under study, with each word frequency capturing the portion of the communication agenda devoted to that word. Using this data set of automated word counts, I calculated the formulas for each candidate measure on the proportion of attention (i.e., number of uses) that each word received relative to the total number of words used in a given month.[11]

The point of Figure 2 is not to see which measure best captures how scope behaved in reality (since I have not yet discussed the substantive nature of the *Times* coverage of the war), but rather to compare the general behavior of these measures over time in a non-simulated environment. Notice that the Herfindahl Index exhibits very little variance. Part of this lack of variance is due to the scaling of Figure 2, but no matter what the scale the Herfindahl Index simply does not respond to whatever real-world signals of scope are driving the other three measures.[12] Notice also that, while they behaved very differently in Figure 1, the two evolutionary factor analysis measures track each other closely in Figure 2.

[11] Thus, each distinct word in this analysis represents a distinct dimension.

[12] I have intentionally excluded the axis labels from Figure 2 because of the very different scales the measures use in the context of real-world data. The differences in how the formulas are calculated means that including the axis labels (four different ones in this case, one for each different measure scaled in a different way) would offer no meaningful information for comparison.

Without knowing anything about the policy discussion at hand, we can make one more critical observation about the measures shown in Figure 2. The two evolutionary factor analysis measures and Shannon's H entropy display rich variance over the time period shown. But the dynamic behavior of the Shannon's H series is quite different from that of the two factor analysis series. While the factor analysis series exhibit periodic deviations (dips or spikes) around what looks to be an equilibrium, the Shannon's H series shows less short-term volatility but much more long-term mobility.

Thus, in considering how scope "really" behaved during this time period, we should look in particular at which of these dynamic behaviors better matches what our theory and a qualitative examination of the *Times* stories tells us about how the scope of the war debate changed over this time period.

To answer this question, we turn to Figure 3. In this figure, we see how the substantive dimensions used in media coverage of the war have shifted over time. This data comes from a project involving the hand-coding of all front-page *Times* stories on the war according to an intricate codebook containing 452 distinct frames that could theoretically be used in discussing the war, all clustered into 12 major dimensions of debate (Boydstun and Glazier 2008). These 12 dimensions are the ones captured in Figure 3, and they are listed in order of their appearance in the chart (top to bottom) in the legend to the right.

The first thing to learn from Figure 3 is that the overall amount of attention to the war—the number of stories on the issue each month, as indicated by the top of the stacked area graph—has changed dramatically over time. In particular, we can observe three major spikes in attention: the first immediately following the terrorist attacks of September 11th and lasting into the first deployment of U.S. troops into Afghanistan in October, 2001; the second during the initial deployment of U.S. troops into Iraq in March, 2003; and the third coinciding with the breaking news in April, 2004 of detainee abuse at Abu Ghraib. Looking more closely at the dimensions used in *Times* stories throughout the course of the war, we see that the debate has changed in important ways over time. Framed at first as an issue of terrorism and then as an issue of weapons of mass destruction, the debate has evolved to focus on very different dimensions, such as prisoner treatment and reconstruction.

Most important for our discussion of scope, we see that while media coverage of the war has contained multiple dimensions throughout these years, a close look reveals that the scope of the coverage has actually increased over time. Granted, this statement is based on an informal examination of the layers of dimensions in Figure 3. But these layers show that the average number of dimensions activated in the debate has increased slightly over time (the slope of the trend line is indeed positive). The number of dimensions has increased even though the average total number of stories written each month (i.e., the number of opportunities dimensions have for being discussed) has declined. Moreover, this increase has corresponded with a diffusion of attention across dimensions. Rather than one or two dimensions consuming the vast bulk of attention, as in the early stages of the war, the latter months show that attention is divided much more evenly.

In terms of how this change has happened, it seems from Figure 3 that media coverage of the war has been highly susceptible to turnover. The component dimensions of the media discussion as well as the scope of the discussion have changed frequently, but despite the large surges in amount of attention given to the war, the scope of discussion has not been quite as volatile. Certainly, it looks like there have been important shifts in the scope of coverage, but even in periods of low overall attention it seems that scope remains relatively high.

Looking back to the behavior of each measure in Figure 2, it is even more apparent that Shannon's H entropy offers the most accurate representation of real-world scope, at least in the context of media coverage of the war. The entropy measure captures both of the major elements of scope suggested in Figure 3: it increases slightly over time and it changes frequently but usually not erratically. As is likely already evident, Figure 3 shows the line graph of the Shannon's H measure of scope calculated from the *Times* word count data superimposed over the stacked area graph showing the substantive dimensions of the discussion. Taking all that we know about scope into account, this line graph is a solid representation of how the latent scope variable has operated in this case.

Automated Coding vs. Hand Coding

Having identified Shannon's H entropy as the best measure of the scope of political communication, both abstractly and in the context of real data, I will briefly discuss the trade-offs between using communication data coded by hand and communication data coded through automated content analysis. The benefits of using hand-coded data are few but significant, chief among these being the ability to stay in control of the data collection process and know personally—usually as the result of countless hours spent reading thousands of texts—what the data set has to say. The main drawbacks to using hand-coded data are: first, the danger of relying on human-developed subjective coding schemes; second, the general infeasibility of tracking attention to *all* dimensions used in a text (not just the primary dimension); and, third, the considerable amount of time and labor that hand-coding requires.

The advantage of being able to stay in control of and connected to one's data is enough to convince most communication scholars to invest the time required for hand-coded data—or abandon the effort altogether. And it is certainly true, at least in my experience, that there can be no complete substitute for hand-coded data. At a minimum, communication scholars should get to know the texts they are studying through examining a systematic sampling of the texts and should re-visit the texts periodically throughout any data collection effort to make sure that the design of the research project continues to reflect the parameters and characteristics of the raw subjects (i.e., the pieces of communication) under consideration. This suggestion is especially true when considering that a complicated text analysis algorithm and a little bit of knowledge can be a dangerous combination.

That being said, my analysis of competing scope measurements suggests—again, consistent with my experience—that automated content analysis can be a worthy tool for researchers to employ. The time that automation can save a researcher is great, but this is actually not the reason I propose incorporating automation into the method for measuring scope. The other two reasons mentioned above for using automation are both more important. First, using automated content analysis removes the subjectivity of the researcher from the design of the coding scheme. Even in the best of circumstances with the clearest of intentions, researchers often approach a research question with preconceived notions about the answer, and these preconceptions can have a viral effect on the structure and content of the resulting data sets.

Take, for example, the coding scheme developed for tracing front-page attention to the war. The coding scheme we developed for this project is about as sophisticated as human coding schemes can get and we were about as careful in developing and vetting this coding scheme as anyone could be. Yet still we had no choice in the end but to force our own understanding of the world onto the texts we were studying. We identified 452 distinct frames comprising 12 distinct dimensions because these were the frames and dimensions that we observed. But another person's observations may have been much different. Why 12 and not 10? 15? 98? Since automated content analysis—at its most basic level—involves taking a count of each word used and then subjecting these word counts to a formula such as Shannon's H entropy, there is no chance that the researcher's perspective will cloud the research design. Even if the researcher wanted to, she could not bias her findings using this method.

Second, employing automated content analysis allows the researcher to investigate *all* the different dimensions used in a given political venue at a given point in time, not just the primary dimension usually captured in hand coding. For example, a data set containing word frequencies for all the words in front-page stories about the war in a given month has a lot more data to offer than does a data set containing a single observation (i.e., the main dimension the human coder identifies in a text). In Figure 4, we see two measures both using Shannon's H entropy: one based on automated content analysis and one based on hand-coded data. Although the formula is the same and the underlying front-page stories are the same, the two measures offer vastly different pictures of how the scope of media attention about the war has unfolded over time.

Which one is more accurate? In addition to the elements I have already discussed in favor of the automated content analysis version of Shannon's H, the weakness of the hand-coded measure (based on some 2,000 stories, by the way) is evident in Figure 4. The low values that plague this series are consistent with a data set that has too few observations. In particular, we can see that the hand-coded measure in Figure 4 tracks strongly with the total amount of front-page attention given to the war seen as the top line of the stacked area graph in Figure 3. Although the amount of attention an issue receives and the scope of that attention are strongly inter-related (Boydstun 2008), still an increase/decrease in overall attention should not necessarily (i.e., mathematically) result in an increase/decrease in scope. By shifting from using only hand-coded data to using automated content analysis in conjunction with hand-coded data, we can take advantage of an enormous stockpile of information contained in political communication texts: words.

Thus, across these many ways of considering the scope of political communication, one method suggests itself as the best approach to measuring the scope of communication: Shannon's H entropy calculated from automated word count data.

Conclusion

The analysis I have offered here is limited in two important ways. First, in order to demonstrate how Shannon's H entropy works on real data, I am relying on a single policy issue (albeit an important one). Thus, when I describe this process as an illustration of the method, I mean just that. After showing how the measure functions in this one case, I hope to have convinced the reader that this measure performs well—in this one case. If my theoretical discussion of scope measures tested against simulated data above was compelling, the reader can infer that Shannon's H entropy is the best measure of scope across contexts, but this conclusion would be an inference only.

Second, and more importantly, because scope is a complex and *latent* variable, we have no way of empirically assessing the relative accuracy of each possible measure. I have made, I hope, a compelling argument for why Shannon's H entropy is a strong theoretical fit for the scope of political communication. But there are no decisive p-values or the like that I can point to in support of my claim. Since our methods should always be driven by our theories—even more than by our p-values—the lack of a verifiable "correct answer" should not deter us. Still, the fact that each measure of scope paints a very different picture is frustrating. Even Shannon's

H entropy measure looks radically different when performed on two different data sets collected *from the exact same news articles*.

Yet the elusive nature of scope as a research variable also reminds us why this line of inquiry is so important. Although scope has already been proven to be a political phenomenon of significant impact, the wide variance in methods used to measure scope puts researchers in this young field of study in jeopardy of talking past one another. We need a system of measuring scope that is as reliable as the variable is important. I propose using Shannon's H information entropy, calculated on data yielded from automated content analysis of communication texts, as the single best method available for measuring scope.

Figure 1. Comparing Different Measures of Scope in the Abstract

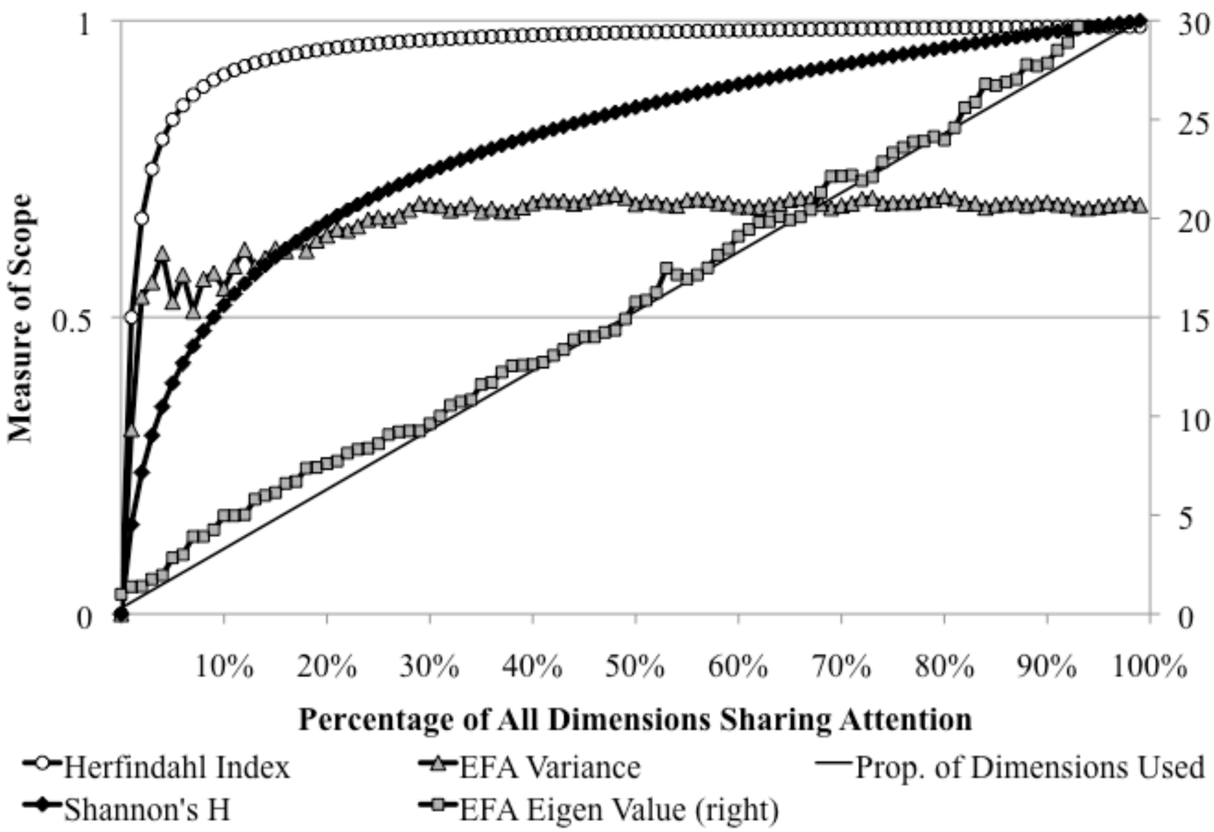


Figure 2. Comparing Measures of Scope in *New York Times* Front-Page Coverage of the U.S. Conflicts in Afghanistan and Iraq, 2001–2005, Using Word Count Data Collected through Automated Content Analysis

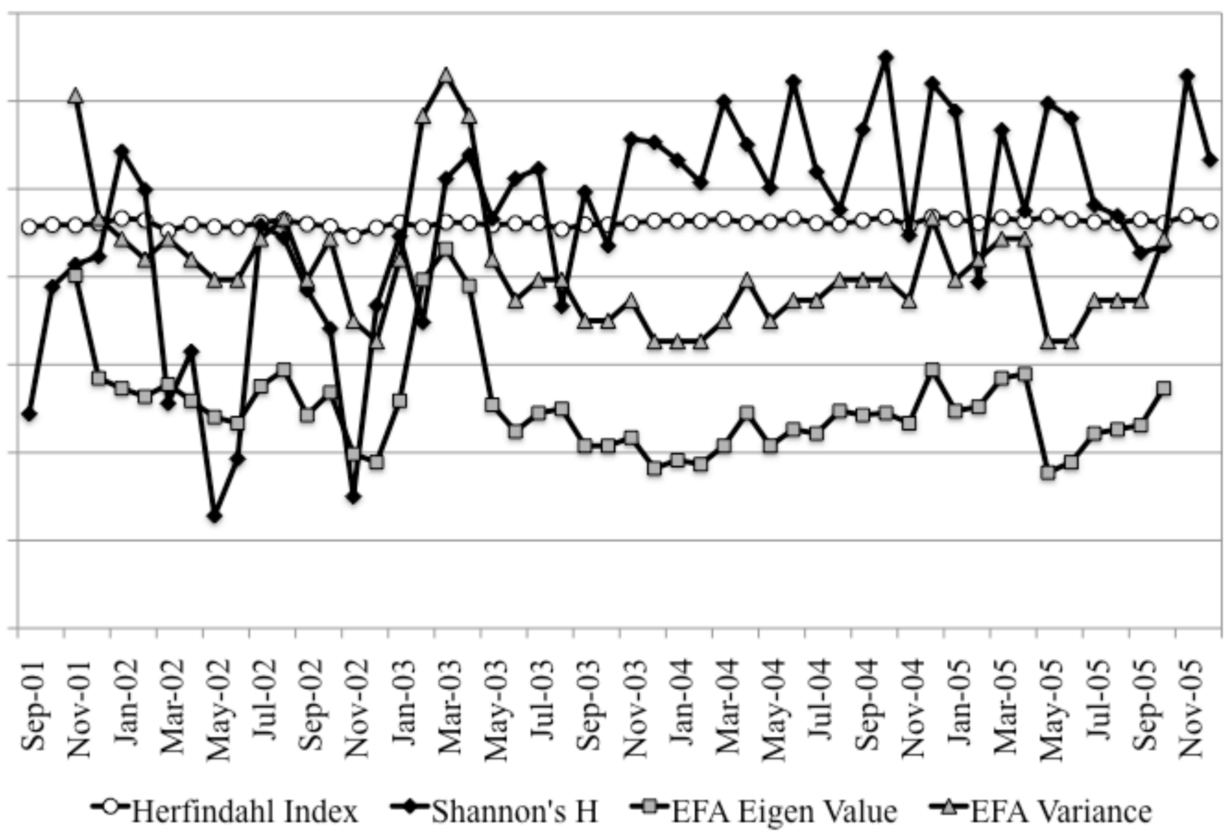


Figure 3. Shannon's H Measure of Scope Based on Automated Content Analysis Data Transposed with Substantive Dimensions of *New York Times* Coverage of the U.S. Conflicts in Afghanistan and Iraq, 2001–2005, Based on Hand-Coded Data

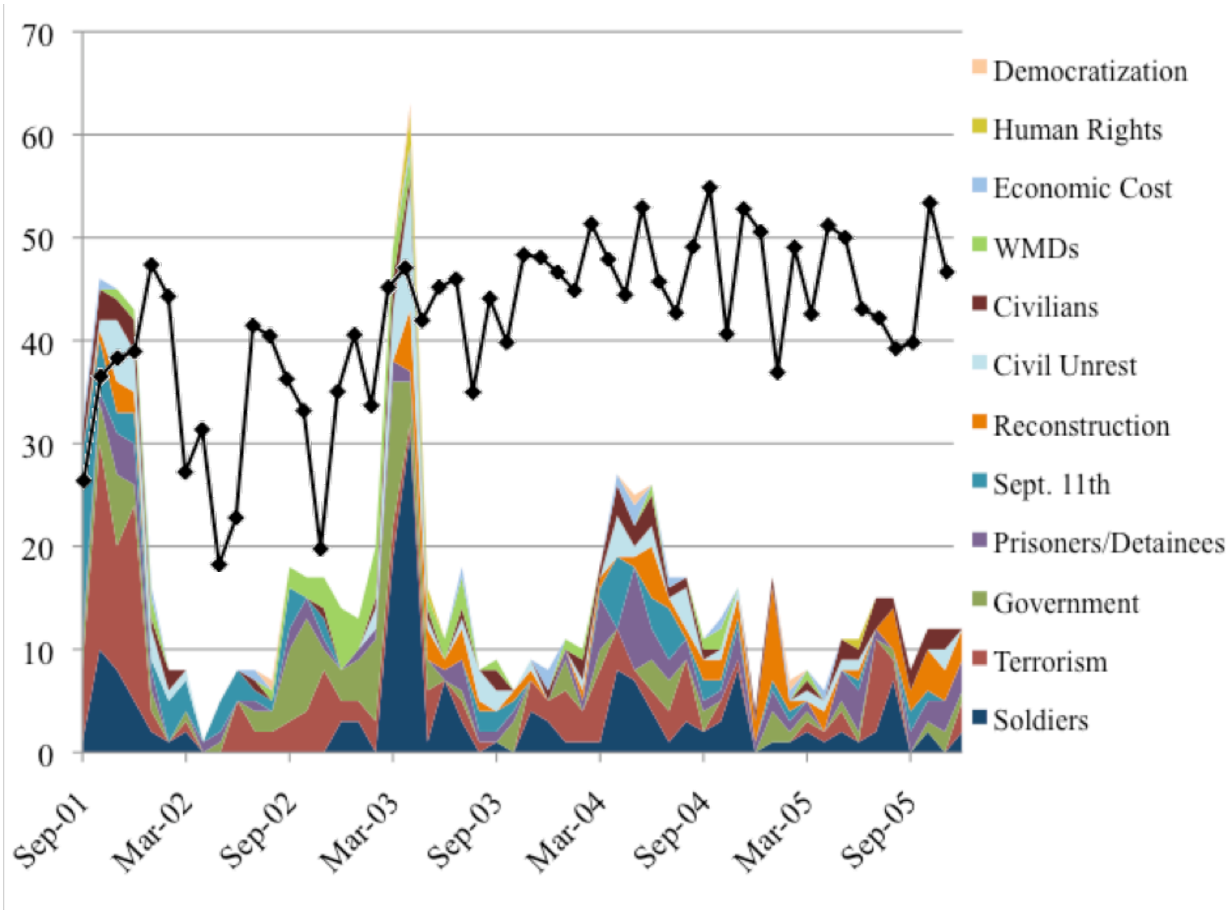
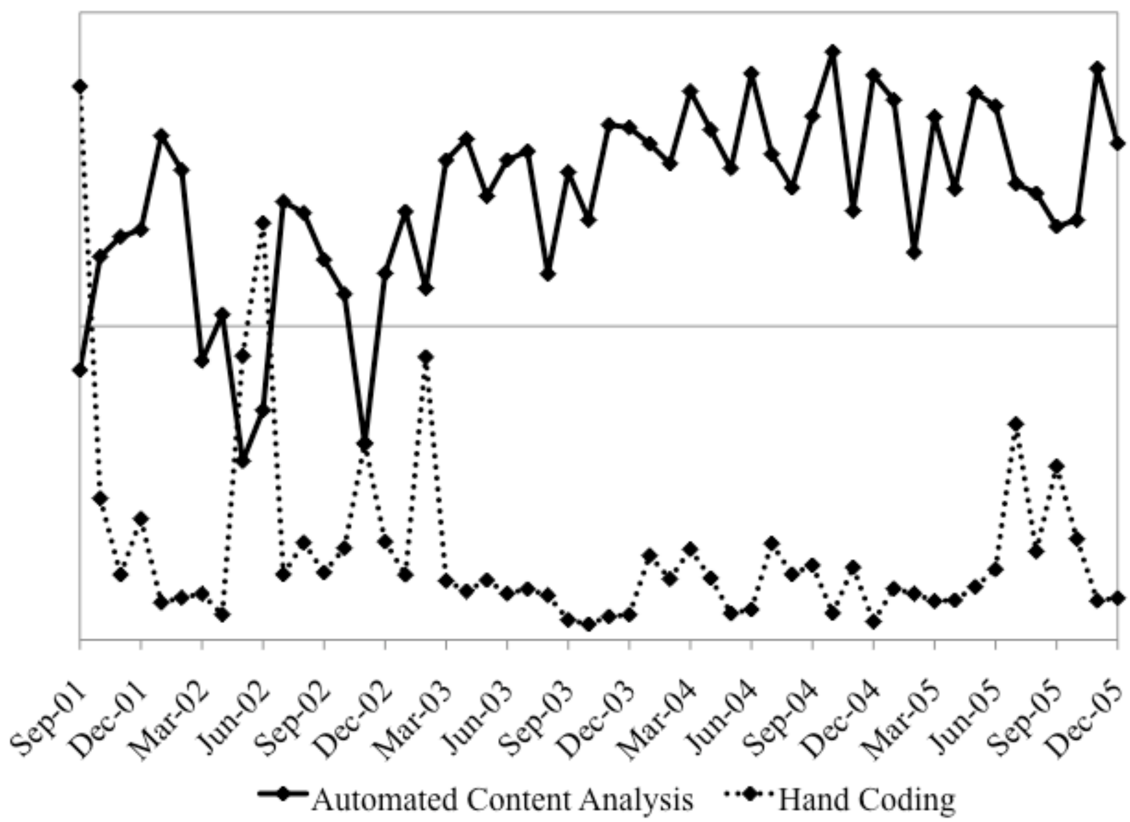


Figure 4. Comparing Shannon's H Entropy Measure of Scope Based on Data Collected through Automated Content Analysis and Data Based on Hand Coding, Both Using *New York Times* Front-Page Coverage of the U.S. Conflicts in Afghanistan and Iraq, 2001–2005



Bibliography

Baumgartner, Frank R., Suzanna De Boef, and Amber E. Boydstun. 2008. *The Decline of the Death Penalty and the Discovery of Innocence*. New York, NY: Cambridge University Press.

Baumgartner, Frank R. and Bryan D. Jones. 1993. *Agendas and Instability in American Politics*. Chicago, IL: The University of Chicago Press.

———. 2002. *Policy Dynamics*. Chicago, IL: University of Chicago Press.

Behr, Roy L. and Shanto Iyengar. 1985. "Television News, Real-World Cues, and Changes in the Public Agenda." *Public Opinion Quarterly* 49 (38): 38-57.

Best, Joel, ed. 1995. *Images as Issues*. New York, NY: Aldine de Gruyter.

Bevan, Shaun. 2008. "Effects on the Scope of Agendas: The Reasons Why the Dog Did Not Bark." Working Paper.

Birkland, Thomas A. 1997. *After Disaster: Agenda Setting, Public Policy, and Focusing Events*. Washington, DC: Georgetown University Press.

Blumer, Herbert. 1948. "Public Opinion and Public Opinion Polling." *American Sociological Review* 13 (5): 542-9.

Boydstun, Amber E. 2008. "How Policy Issues Become Front-Page News." Doctoral Dissertation, The Pennsylvania State University.

Boydstun, Amber E. and Rebecca Glazier. 2008. "From Spreading Freedom to WMDs and Back Again: Framing Dynamics and the War on Terror." Paper resented at the Midwest Political Science Association Annual Meeting, April 3-6, 2008 in Chicago, IL.

Brewer, Paul R. and Kimberly Gross. 2005. "Values, Frames, and Citizens' Thoughts about Policy Issues: Effects on Content and Quality." *Political Psychology* 26 (6): 929–48.

Callaghan, Karen J. and Frauke K. Schnell, eds. 2005. *Framing American Politics*. Pittsburgh: University of Pittsburgh Press.

Chong, Dennis and James N. Druckman. 2007. "A Theory of Framing and Opinion Formation in Competitive Elite Environments." *Journal of Communication* 57: 99-118.

Cobb, Roger W. and Charles D. Elder. 1972. *Participation in American Politics: The Dynamics of Agenda-Building*. Baltimore, MD: Johns Hopkins Press.

Dalton, Russel, J., Paul Allen Beck, Robert Huckfeldt, and William Koetzle. 1998. "A Test of Media-Centered Agenda Setting: Newspaper Content and Public Interests in a Presidential Election." *Political Communication* 15 (4): 463-81.

Dardis, Frank E., Frank R. Baumgartner, Amber E. Boydstun, Suzanna De Boef, and Fuyuan Shen. 2008. “Media Framing of Capital Punishment and Its Impact on Individuals' Cognitive Responses.” *Mass Communication and Society* 11 (2): 115-140.

Dearing, James W. and Everett M. Rogers. 1996. *Agenda-Setting*. Thousand Oaks, CA: Sage Publications.

Edwards, George C. and B. Dan Wood. 1999. “Who Influences Whom? The President, Congress, and the Media.” *The American Political Science Review* 93: 327-44.

Entman, Robert M. 1989. “How the Media Affect What People Think: An Information Processing Approach.” *The Journal of Politics* 51 (2): 347-70.

Gamson, William A. and Gadi Wolfsfeld. 1993. “Movements and Media as Interacting Sysmtes.” *The ANNALS of the American Academy of Political and Social Science II* 528: 114-125.

Gilliam, Franklin D., Jr. and Shanto Iyengar. 2005. “Super-Predators or Victims of Societal Neglect? Framing Effects in Juvenile Crime Coverage.” In *Framing of American Politics*, ed. K. J. Callaghan and F. K. Schell. Pittsburgh: University of Pittsburgh Press.

Hilgartner, Steven and Charles Bosk. 1988. “The Rise and Fall of Social Problems: A Public Arenas Model.” *The American Journal of Sociology* 94 (1): 53-78.

Iyengar, Shanto. 1991. *Is Anyone Responsible? How Television Frames Political Issues*. Chicago, IL: University of Chicago Press.

Iyengar, Shanto, and Donald R. Kinder. 1987. *News That Matters: Television and American Opinion*. Chicago, IL: University of Chicago Press.

Iyengar, Shanto, and Jennifer A. McGrady. 2007. *Media Politics: A Citizen’s Guide*. New York, NY: W.W. Norton & Company.

Jacoby, William G. 2000. “Issue Framing and Public Opinion on Government Spending.” *American Journal of Political Science* 44 (4): 750-67.

Jasperson, Amy E., Dhavan V. Shah, Mark Watts, Ronald J. Faber, and David P. Fan. 1998. “Framing and the Public Agenda: Media Effects on the Importance of the Federal Budget Deficit.” *Political Communication* 15 (2): 205-24.

Jerit, Jennifer. 2008. “Issue Framing and Engagement: Rhetorical Strategy in Public Policy Debates.” *Political Behavior* 30 (1): 1-24.

Jones, Bryan D. and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems.* Chicago, IL: University of Chicago Press.

Kellstedt, Paul M. 2003. *The Mass Media and the Dynamics of American Racial Attitudes*. New York: Cambridge University Press.

Kinder, Donald R. and Thomas E. Nelson. 2005. "The Democratic Importance of Elite Frames." In *Framing American Politics*, ed. K. J. Callaghan and F. K. Schell. Pittsburgh: University of Pittsburgh Press.

Kingdon, John W. 1995. *Agendas, Alternatives, and Public Policies*. New York, NY: HarperCollins.

Landauer, Thomas K. and Susan T. Dumais. 1997. "A Solution to Plato's Problem: The Latent Semantic Analysis Theory of the Acquisition, Induction, and Representation of Knowledge." *Psychological Review* 104: 211-240.

Laver, Michael, Kenneth Benoit, and John Garry. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97 (2): 311-331.

McCombs, Maxwell. 2004. *Setting the Agenda: The Mass Media and Public Opinion*. Cambridge, UK: Polity Press.

McCombs, Maxwell and Salma I. Ghanem. 2003. "The Convergence of Agenda Setting and Framing." In Reese, Stephen D., Oscar H. Gandy, Jr. and August E. Grant, Eds. *Framing Public Life*. Mahwah, NJ: Lawrence Erlbaum Associates.

Nelson, Thomas E., Rosalee A. Clawson, and Zoe Oxley. 1997. "Media Framing of a Civil Liberties Controvery and its Effect on Tolerance." *American Political Science Review* 91 (3): 567-84.

Nelson, Thomas E. and Zoe M. Oxley. 1999. "Framing Effects on Belief Importance and Opinion." *Journal of Politics* 61 (4): 1040-67.

Pollock, Philip H. III. 1994. "Issues, Values, and Critical Moments: Did "Magic" Johnson Transform Public Opinion on AIDS?" *American Journal of Political Science* 38 (2): 426-46.

Riker, William H. 1986. *The Art of Political Manipulation*. New Haven, CT: Yale University Press.

Schattschneider, E.E. 1960. *The Semi-Sovereign People*. New York: Holt, Rinehart, and Winston.

Sheingate, Adam D. 2006. "Structure and Opportunity: Committee Jurisdiction and Issue Attention in Congress." *American Journal of Political Science* 50 (4): 844-859.

Snow, David A., Sarah A. Soule, and Hanspeter Kriesi. 2004. *The Blackwell Companion to Social Movements.* Oxford, England: Blackwell.

Spector, Malcolm and John I. Kitsuse. 1973. "Social Problems: A Reformulation." *Social Problems* 21 (2): 145-59.

———. 1977. *Constructing Social Problems*. Hawthorne, NY: Aldine de Gruyter.

Terkildsen, Nayda and Frauke K. Schnell. 1997. "How Media Frames Move Public Opinion: An Analysis of the Women's Movement." *Political Research Quarterly* 50 (4): 879-900.

Walgrave, Stefaan and Peter Van Aelst. 2006. "The Contingency of the Mass Media's Political Agenda Setting Power: Towards a Preliminary Theory." *Journal of Communication* 56 (1): 88-109.

Wolfe, Michelle. 2008. "Public Opinion Signaling and Signal Strength: Issue Priority, Issue Salience and Information Environments." Working Paper.

Wood, B. Dan and Alesha Doan. 2003. "The Politics of Problem Definition: A Theory and Application to Sexual Harassment." *American Journal of Political Science* 47 (4): 640-53.

Wood, B. Dan and Jeffrey S. Peake. 1998. "The Dynamics of Foreign Policy Agenda Setting." *The American Political Science Review* 92 (1): 173-84.