

Abstract

In authoritarian states, mass mobilization is rare, as information asymmetries are the norm. Both regime actors and activists are uncertain as to the distribution of support for their respective causes and the likelihood of mobilization. Activists use the media strategically to persuade potential allies and to gauge the success of their tactics. Regimes, likewise, try to deter mobilization and reassure allies. Bystanders that are neither activists nor regime allies are the prize in a battle for support. I argue that contests over dominant narratives in the media represent the openings of political opportunities and can establish common knowledge for bystanders unsure about the costs and benefits of participating in collective action. Using more than 29,000 Arabic news articles from 2010 and 2011, I employ a novel method of quantitative text analysis to demonstrate that the Mubarak regime attempted to demobilize activists by emphasizing themes of danger and instability, highlighting its own actions to reassure potential allies, and attempting to minimize coverage of successful mobilization. Conversely, independent and international media downplayed danger and instability, largely ignore the regime's statements, and focused on successful mobilization. I contend that the Mubarak regime lost control of the dominant narrative before protests even began, signaling opportunity for mobilization.

The Battle for Bystanders: Information, Meaning Contests, and Collective Action in the Egyptian Uprising of 2011

Large-scale collective action under authoritarianism is often unexpected (see, e.g., Kuran, 1991, 1997). Anti-regime mobilization in non-democracies is costly and dangerous; state security apparatuses and information asymmetries can make success seem unlikely or impossible (Lichbach, 1994). And yet, throughout history, individuals, often bystanders with little prior political involvement, have taken to the streets to oppose autocracy. Most recently, a wave of uprisings swept and continues to sweep across the Middle East and North Africa. The largest of these uprisings took place in Egypt over just 18 days in 2011, involving hundreds of thousands of protesters and the eventual resignation of Hosni Mubarak, the Egyptian President for 29 years.

Roughly two months before the uprising, many of Mubarak's allies secured their political offices in elections widely seen as among the country's most fraudulent ever. Although popular grievances over the elections were simmering and the economy continued to under-perform, mass protests did not materialize. The regime surely must have felt confident about its position and stability. Yet on 25 January 2011, thousands mobilized against the Mubarak regime in Alexandria, Mansura, Aswan, and other cities. In Cairo, protesters gathered in Tahrir (Liberation) Square--the name of which would become forever associated with the ouster of one of the world's longest-sitting autocrats. Once seemingly immovable, Mubarak ceded the presidency on 11 February 2011. Given that grievances were apparently constant and the police state had not relaxed its grip, what explains the timing of these protests? Why did individuals feel that collective action was possible in January but not in the preceding months when the regime flaunted its corruption?

While we may never know what was going through the minds of the protesters, it is possible to make inferences about salient information available to protesters, bystanders, and the regime. The role of the media in protest is well-documented. Activists use the media strategically to persuade potential allies (see, e.g., Benford and Snow, 2000; Snow, 2007) and to gauge the success of their tactics (Koopmans and Olzak, 2004). Regimes, like movement actors, use the media strategically to

deter mobilization and reassure allies. The all-important mass of bystanders that are neither activists nor regime allies are the prize in a battle for support. As both movement actors and regimes have an information deficit with respect to the political loyalties of the bystanders, each must attempt to persuade bystanders to either mobilize or remain at home. The mass media are a central arena for this contest (Gamson, 2007).

I argue that contests over dominant narratives in the media represent the openings of political opportunities and can establish common knowledge for bystanders unsure about the costs and benefits of participating in collective action (Coyne and Leeson, 2009). The participation (or non-participation) of non-committed bystanders is crucial for the success or failure of mass mobilization. These bystanders, who mostly lack the ideological or instrumental commitment to either the regime or the opposition, must be won over. To do so, both pro- and anti-regime groups must engage in a contest over the interpretation of political reality -- the battle for bystanders (Gamson, 2007).

Clearly, I do not claim that the media caused a revolution in Egypt. Nor do I contend that the mass media solve all of the coordination problems necessary to mobilize. However, I argue that the mass media are a frequently overlooked source of political information--even in authoritarian societies where the media are controlled. As a large literature in social movements studies outlines, content and framing are important for mobilization (Koopmans and Olzak, 2004; Polletta, 2008; Snow, 2007). However, it is not only activists who are using the media to persuade bystanders; states engage in the same behavior. This process is *dynamic* and *strategic* -- pro- and anti-regime media outlets watch one another as well as events on the ground to determine the best way to win bystander support. In addition to the recognized importance of the amount and tone of media coverage, the interplay of coverage both within and between media sources conveys information about the regime and the opposition.

Using the case of the Egyptian uprising of 2011, I argue that the Mubarak regime attempted to signal that a political opportunity for collective action did not exist and that it would be dangerous and unnecessary; it did so by emphasizing the danger and destabilizing character of protest while simultaneously trying to signal to supporters and fence-sitters that mobilization was unnecessary

to achieve reforms. Meanwhile, as the regime downplayed coverage of opposition actors and successful mobilization, independent and international media adopted unified narratives that contradicted the regime's version of events, suggesting the possibility and desirability of mass mobilization to achieve political change.

The message that the regime tried to transmit was not received successfully and was unable to counteract the effects of opposition messages. A week into the uprising and as protesters continued to number in the hundreds of thousands, the state media abandoned its attempt to control the narrative and began to resemble independent media. The inability of the regime to impose its interpretation of political reality offered signals that space for opposition had opened; the costs of mobilizing were falling and the benefits were rising.

To do so, I use full-text news articles from 1 November 2010 until 28 February 2011 and a novel method of quantitative text analysis to model the content of the mass media before, during, and after the 18-day uprising. Media data provide day-by-day snapshots of what contemporary actors were seeing and discussing; these data do not suffer from retrospective bias and are unaffected by post hoc narratives attached to the events. The data represent the views of the regime, the independent media, and international media. News articles are purposively generated--the choice of words and topics are instrumental and intended to inform and persuade. State media employees were overseen by the Egyptian Ministry of Information and thus state media are regime actors. In addition to representing the events as contemporary actors understood them, media data provide traces of regime and opposition strategic action.

Background: Arab Uprisings of 2010-2011

The history of the Arab Uprisings¹ of 2010-2011 is still being written. For non-area experts, I provide a thumbnail sketch of events.

¹The name for the protests that unseated Ben Ali, Mubarak, Gaddafi, and Saleh is disputed. Western media tend to use "Arab Spring" whereas Arabic media tend to refer to the "Arab Revolutions" and "Arab Awakening." I use the neutral term "uprisings" to avoid entering into normative debates about nomenclature and because it is unclear which of the events qualify as revolutions.

On 18 December 2010, a street vendor named Mohamed Bouazizi sets himself on fire to protest state and police injustice; he later succumbed to his wounds on 4 January 2011. Following his self-immolation, protests begin in the small hometown of Bouazizi, Sidi Bouzid, and spread town-by-town to the capital, Tunis. The protests culminate in President Zine El Abidine Ben Ali, who ruled for more than 23 years, fleeing to Saudi Arabia on 14 January 2011. While protests grew in Tunisia, similar protests and riots erupted in Algeria, Libya, and Yemen. A single protester self-immolates in front of the Egyptian Parliament as well.

As restive populations began to mobilize across the Middle East and North Africa, Egyptian activists call for mass protests on 25 January 2011 -- Police Day, a national holiday. The use of pro-government national holidays to stage protests is a common tactic in autocratic states, as popular mobilization on these days is permitted. The activists who call for the protests include the *Kefaya* (Enough!) movement, the 6 April Movement, and a Facebook group entitled "We Are All Khaled Saeed" later revealed to be administered by Google executive Wael Ghonim. However, the largest opposition group in Egypt, the Muslim Brotherhood, is ambivalent about participating in the protests and only sanctions them at the last minute. Coptic Christians are discouraged from participating by church officials, following a deadly sectarian attack on a Coptic church in Alexandria on New Year's.

Much larger than expected protests take place in multiple Egyptian cities on 25 January. Emboldened by the unexpected turnout, further protests are called for Friday, 28 January, dubbed the "Day of Rage" or "Day of Anger." Although some smaller protests continue on the 26th and 27th, hundreds of thousands return to the streets on Friday. This cycle of protest, a mid-week protest followed by a protest on Friday, continues throughout the uprising. At this point, fearing further protests, the government effectively disconnected Egypt from the internet and restricted most cellular communications for the next several days.

Mubarak announced limited concessions, sacking his cabinet, but is mostly critical of the protests in a national address and scolds the protesters to leave public squares. The following day, Omar Suleiman is named vice president and a new prime minister, Ahmed Shafiq is named. Egypt had

been without a vice president (and thus without a clear successor to the Presidency) for Mubarak's entire 29-year rule. Another large protest materializes on the 31st of January; encouraged by the continuing mobilization, activists call for a "March of Millions" the following day, 1 February. Although numbers vary wildly, most sources agree that at least 250,000 took to the streets.

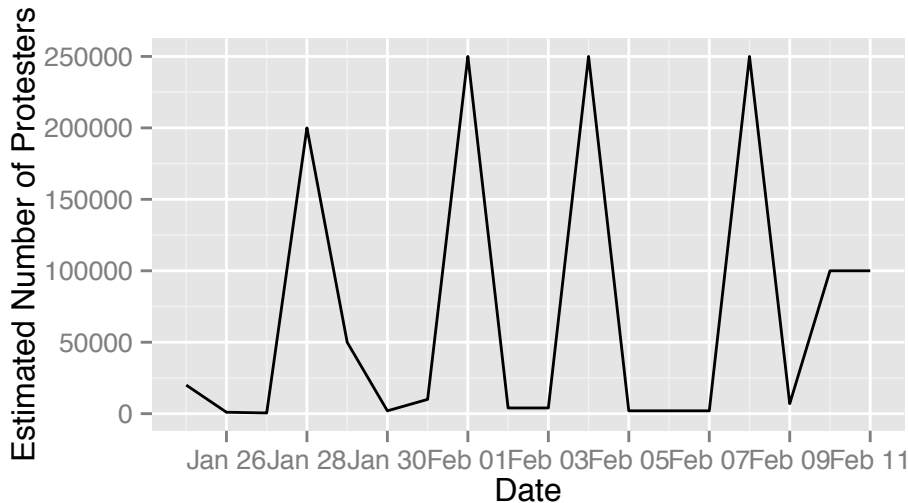
Following this huge turnout on 1 February, pro-Mubarak protesters flooded Tahrir Square and a pitched battle involving knives, bats, camels, and horses took place in Tahrir over the next two days. Undeterred, protesters returned *en masse* to Tahrir on Friday, 4 February for what is dubbed the "Day of Departure." The next day, the leadership of Mubarak's political party, including Mubarak's son and assumed successor Gamal, resigned their posts. The cycle repeats and huge crowds returned to Tahrir on the 8th following two days of smaller protests. Constitutional reforms are announced and the government acts to prevent a bank run. Mubarak orders a review of the fraudulent elections of 2010. Enraged by Mubarak's unwillingness to step down, protests continue to build, ending on 11 February 2011 with Mubarak ceding the powers of the presidency to the Supreme Council of the Armed Forces via a pre-written message read by Vice President Omar Suleiman.

Figure 1 contains extremely rough estimates of protest numbers compiled from the *New York Times*, *Al Jazeera* English, and BBC coverage of the protests. The cyclical pattern of mid-week protest followed by Friday protest is readily observable. These numbers should be interpreted with an excess of caution, as reliable and unbiased estimates of protest turnout are non-existent. Media accounts of turnout were coded essentially on a logarithmic scale, as turnout estimates were usually given in terms of "hundreds", "thousands", "tens of thousands", "hundreds of thousands", and so on. "Thousands" was coded as 2,000 and "tens of thousands" was coded as 20,000. When more specific figures were provided, they were used.

Information, collective action, and agenda-setting

Information and its often unequal distribution figure prominently in contentious politics. As both movement and counter-movement actors wish to inform and mobilize potential allies while demobilizing potential opponents (Lichbach, 1998; Polletta, 2008). Political opportunities must be *perceived* to

Figure 1: Rough estimate of protest numbers for 18 days of protest in Cairo.



be opportunities (Gamson and Meyer, 1996; Koopmans and Olzak, 2004; Kurzman, 1996; Maher, 2010; Meyer, 2004; Roscigno and Danaher, 2001) -- indeed as Kurzman (1996) has argued perceptions of opportunities may be more important than objective opportunities. The media is one of the central arenas for this contest (Andrews and Biggs, 2006; Gamson, Croteau, Hoynes, and Sasson, 1992; Gamson, 2007; Schudson, 2002). When the counter-movement is the state, as it usually is, this contest is necessarily characterized by power asymmetries; even in democracies, state actors have routinized access to the media (Oliver and Maney, 2000; Schudson, 2002) and are more likely to be perceived as neutral. Additionally, media attention helps establish common knowledge--knowledge that a large number of individuals possess as well as the knowledge that other individuals are likely to hold this knowledge (Chwe, 1999).

However, in non-democracies, the state often controls, partially or totally, large parts of the media. This exacerbates information asymmetries and perpetuates pluralistic ignorance, a state in which individuals are unaware of the true beliefs of their peers. Preference falsification and the pluralistic ignorance that it creates can sustain unpopular authoritarian regimes and suppress

collective action (Kuran, 1997; Lichbach, 1998; Lohmann, 1994). However in all but the most tightly controlled media environments, exogenous events can create instability in the distribution of information (Pfaff, 2006). As access to information increases and the costs of doing so decrease, authoritarian regimes must become increasingly agile in managing information and its interpretation (Lynch, 2011).

The media cover events deemed newsworthy; however, as has been pointed out in many previous works, they are not impartial arbiters in this process (Andrews and Caren, 2010; Oliver and Meyer, 1999; Oliver and Maney, 2000; Roscigno and Danaher, 2001; Schudson, 2002). Thus, while one would expect that different media outlets would cover the same events, decisions such as the content, framing, and quantity of the coverage affect how consumers of the media interpret those events. The media environment is characterized by a finite amount of space and the attentional capacity of the public is also finite. The literature on "news values" and "news-making" is a well-developed one, focusing on how media actors make decisions about what to cover as well as how and when to cover it (for an overview, see Schudson (2002)).

These decisions have important agenda-setting effects and influence individual perceptions of what issues are important as well as what aspects of those issues are important (Amenta, Caren, Olasky, and Stobaugh, 2009; Andrews and Caren, 2010; Brosius and Weimann, 1996; Larcinese, Puglisi, and Snyder Jr, 2011; McCombs and Shaw, 1972, 1993; Scheufele, 1999). In non-democracies, these decisions are constrained by either explicit or self-imposed censorship (Zhao, 1998), biasing media agendas toward regime preferences. However, as the agenda-setting power of the media may be attenuated when the media are viewed as untrustworthy (Tsfati, 2003), the relationship between the media agendas and public agendas in non-democracies is not as straightforward as in democracies. In these cases, the interplay between regime-friendly and independent² media may be as important as the attention paid by each outlet to various issues.

Regime control of the media was simpler when informational borders were less porous and mass

²I use the terms independent and non-state media interchangeably in this piece. To my knowledge, there are currently no data that rank or quantify the relative independence of media outlets besides impressionistic "state-owned" vs "non-state-owned."

media outlets were relatively few in number. Manipulating information is a long-standing practice by authoritarian regimes, but the number of hermit states such as North Korea is declining. As an emerging literature demonstrates, modern authoritarianism is diverse, relying on a combination of institutional and electoral manipulation (Blaydes, 2006; Brownlee, 2007; Levitsky and Way, 2010; Slater, 2009). As Koopmans and Olzak (2004) argues, as the number of channels of communication increases, more information can enter into the media environment. This does not solve the problem of individual attentional capacity, and, in fact, may exacerbate attempts by both movement and counter-movement actors to make their message highly visible, resonant, and legitimate. Additional communicative outlets afford expanded opportunities for actors to self-select into media consumption. If this is the case, ideological members of the polity are likely to self-select into media with coincident ideological leanings; in doing so, the role of non-committed bystanders increases in importance. Media outlets that cater to specific segments of the polity convey less information on average than non-partisan outlets in that they are more predictable in their coverage than less partisan outlets. By consuming media from partisan outlets, the non-committed center of the polity is therefore unlikely to receive novel information about the size of regime loyalists and the opposition; they are also unlikely to receive novel information about the likelihood of success of anti-regime collective action if the signal received from both outlets is relatively constant.

However, changes in the *distribution* of pro- and anti-regime media signal space for meaning contests for bystander allegiances. As the regime holds the upper hand in controlling the amount of and freedom of independent media, the regime's legitimating narrative is the *de facto* narrative for bystanders. If the regime cedes narrative control either intentionally or unintentionally, this signals that movement actors have found the requisite space to promote their interpretation of political reality. The counterpart to this, in which independent outlets' narratives align with state media, is a sign of two things -- events are not politically contested or censorship, either explicit or self-imposed.

The battle for bystanders

Assume that an authoritarian polity is composed of three groups: regime loyalists, the opposition, and non-committed bystanders. Each of these groups are composed of heterogeneous individuals with a distribution of preferences and motivations for their own political alignment. An individual might be allied with the regime for ideological reasons, for fear of persecution, for economic or patronage reasons, and so on. Each of these committed groups has individuals at one tail that are highly committed to the cause and are, at least, highly unlikely to have their allegiances swayed and, at most, more likely to engage in collective action on behalf of their allegiance. Each of these groups also has individuals at the opposite tail of the distribution, who are more likely to drift in and out of the non-committed center depending on perceived incentives and the shifting political landscape. The non-committed mass of bystanders is the largest and most heterogeneous bloc of the polity; Lohmann (1994) has conducted the most complete analysis of the various subgroups that make up the center. The participation of these bystanders is crucial for the revelation of the distribution of political preferences (Lohmann, 1994; Kuran, 1997).

This thus implies that the center is composed of information-impooverished (either purposefully or not) individuals who rely on external cues to estimate the likely composition of the two opposing blocs. However, this does not imply that the center is merely a sponge waiting to absorb information from the opposition or the regime. Rather, much of the center, as discussed by Lohmann (1994), intentionally chooses not to engage in pro- or anti-regime collective action due to significant uncertainty about the dangers and benefits of doing so. In the event of potential shifts in political opportunities, the capture of these bystanders is crucial to success and, as threshold models of collective action explicate (Granovetter, 1978; Lohmann, 1994; Macy, 1991; Oliver and Marwell, 2001).

Authoritarian regimes hold significant advantage in the control of information and its presentation through state-friendly media outlets and through censorship. Therefore, in the absence of major collective action or exogenous events, non-committed bystanders are unlikely to perceive true changes in regime weakness or opposition strength or to incorrectly perceive the presence of opportunities (Kurzman, 1996; Maher, 2010; Sawyers and Meyer, 1999). Regime survival is an equilibrium

state and bystanders perceive collective action against the regime to be undesirable or dangerous, the opposition to be unlikely to succeed in anti-regime collective action, or some combination of both. Gamson (2007) further argues that bystanders are biased toward inaction due to a preference for stability and disinterest in the balance of political power, an argument supported by a large psychological literature on status-quo bias (see, e.g., Kahneman, Knetsch, and Thaler, 1991).

Hence, major reasons why press outlets are tightly controlled in authoritarian states are not only to provide a source of regime-friendly information, perpetuate the regime's legitimating narrative (Kurzman, 1996), and to suppress anti-regime information, but also to signal the ability of the regime to exert its influence over the information environment. Regardless of whether or not the information presented in the regime-friendly media is perceived to be true, the ability to control the presentation of information is an important signal of regime capacity. However, cowing the opposition or demonstrating power to the center is not the sole purpose of controlling the media; an equally important goal is to reassure and rally the loyalist base.

These goals are in tension with one another; pro-regime messages that try to convince the polity that anti-regime collective action will be unsuccessful may galvanize opposition supporters. Further, if regime messages are in stark contrast to events "on the ground", they are likely to stretch the "elasticity of reality" (Baum and Groeling, 2008) -- the regime does not have unlimited space within which to spin events except perhaps in extremely information-impooverished environments (which are rapidly disappearing from the planet). Non-committed bystanders may view attempted narratives that stretch credible interpretations of events as signs of desperation or disconnect with reality. Both attention and inattention to events carry information and the ways in which they are covered across sources with various political leanings also carry information.

While the majority of the framing literature focuses on efforts by movement actors to increase mobilization, it largely neglects the state's counter-mobilization framing efforts. While social movements and mobilization may have become routinized parts of politics in many democratic societies, mass mobilization poses substantial risks for non-democratic regimes. Thus, the regime is motivated to present a narrative that minimizes the perception of political opportunities and success, or maximizes

the costs associated with acting on perceived opportunities. All else being equal, bystanders face a choice between which narrative is more compelling. Thus, bystanders toward the pro-regime end of the distribution will likely give more credence to the regime's narrative and bystanders closer to the anti-regime tail will likely give more credence to non-state media.

However, when the narrative shifts to one direction or the other, bystanders are presented with additional information about the relative position of each side -- either the regime is able to reassert control over the narrative or the opposition's narrative becomes dominant. If the regime cedes the narrative in the face of continued mobilization, a strong signal is transmitted that the opposition has an increased likelihood of success. Note however, that media coverage is only one source of information; it is entirely possible that the regime will maintain or reassert its control over the narrative in the face of continued mobilization.

The regime can achieve its goals in three ways: a) *raise the perceived costs* of participation by emphasizing danger and instability, b) *lower the perceived benefits* of participation by presenting evidence that reform is already occurring or will occur without the need for mobilization, and c) *invalidate perceptions of opportunity* by minimizing coverage of successful mobilization.

Conversely, independent media should present a narrative that increases perceptions of political opportunities and success. These media will a) *lower the perceived costs* of participation by minimizing coverage by deemphasizing danger and instability, b) *raise the perceived benefits* of participation by minimizing evidence of reform occurring without mobilization, and c) *validate perceptions of opportunity* by emphasizing coverage of successful mobilization.

These propositions suggest the following hypotheses:

1. When faced with the threat of mass protests, state media should devote more daily coverage to topics that contain words related to danger and instability than independent or international media.
2. When faced with the threat of mass protests, state media should devote more daily coverage to topics that contain words related to regime actions than independent or international media.

3. When mass protests are successfully held, state media should devote less daily coverage to topics that contain words related to opposition and mobilization than independent or international media.

Methods

Media content are an unobtrusive source of intentionally created data from the perspective of contemporary actors. The use of media data provides a proxy for how both regime and opposition understood events as they unfolded and provide traces of their actions in response to these understandings. Content analysis is the traditional method for analyzing unstructured text; however, the costs to analyze the large number of documents under consideration were prohibitive, especially given that the articles are in Arabic. Therefore, I employ an unsupervised method of text analysis known as topic modeling developed, from machine learning (Steyvers and Griffiths, 2007). Rather than giving each article a close reading using a pre-established coding scheme, I adopt a "text as data" approach (Schrodt, 2010) and algorithmically model the latent mixture of topics that were used to generate the observed documents. Latent Dirichlet Allocation (LDA; (Blei, Ng, and Jordan, 2003; Blei, 2011)) is a Bayesian probabilistic topic modeling procedure that allows for the modeling of large corpora with no prior assumptions except for the number of expected topics and two hyperparameters, discussed below. Prior to estimating the model, the researcher need not know the contents of the documents. Using the observed words in documents, LDA allows one to estimate the overall distribution of topics within the corpus as well as the distribution of topics within documents. Clearly, interpretation and validation of the model output requires substantive expertise, but the front-end costs of analyzing large numbers of documents are diminished substantially using this method.

All documents in a corpus share the same topic distribution, but the per-document topic assignment consists of varying proportions of topics (including $p = 0$, the modal proportion on a per-document basis). The topics themselves are categorical distributions over words in the vocabulary (the set of all words that appear in any document) where words have an associated probability of belonging

to that topic. Words with higher probabilities are more likely to appear when that topic is used to generate (part of) a document; words with lower probabilities are less likely to be used in communicating about a topic. Words can be associated with more than one topic, meaning that LDA is capable of dealing with polysemy and synonymy. This is contrasted with earlier methods of topic modeling that required that words be assigned to a single topic (Blei et al., 2003; Blei, 2011). LDA has been demonstrated to produce human-recognizable and interpretable (Chang, Boyd-Graber, Gerrish, Wang, and Blei, 2009) topics on corpora as diverse as US Senators' press releases (Grimmer, 2010), conference proceedings and academic journals (Hall, Jurafsky, and Manning, 2008), and eighteenth-century newspapers (Newman, 2006).

Note that many of the discovered topics will be irrelevant to collective action. Although this may lead to disregarding a large portion of the articles for this analysis, this approach provides several advantages. First, it allows articles that may not explicitly contain protest-related words to become associated with protest-related topics; these articles would be omitted from a keyword-based sampling strategy. Second, it allows for measuring how much coverage is devoted to particular topics by including all topics found in the source and affords modeling changes in this over time. Third, despite using a large number of articles for analysis, the modeling strategy and inclusive data collection strategy actually decrease human costs of selecting and reading articles for relevance.

Pre-processing

In order to prepare the articles for analysis, some pre-processing is required. The implementation of LDA utilized here is based on a "bag-of-words" model of text analysis; the order of words within documents and the ordering of documents within the corpus is assumed to be unimportant, although the dates of the documents are retained as meta-data for later incorporation into the analysis. While this may seem to be an unrealistic assumption, word order has been demonstrated to be unimportant in many applications (Hopkins and King, 2010). The articles are processed to remove punctuation, numbers, and extra whitespace. Then, the words within documents were normalized as much as possible so that plurals and singulars would be treated as the same word. Similarly, common

prefixes and suffixes were removed³.

Documents are then converted into vectors of word frequencies and combined into a matrix such that rows are documents and columns are the individual words that appear at least once in the entire corpus of documents. The cells of the matrix are the word frequencies of word w within document d . The modal cell value is 0. For the Arabic data, this produced a 29,985 document x 189,596 word matrix. The word frequencies are then transformed via a term-frequency inverse-document-frequency (tf-idf) operation (Salton, Fox, and Wu, 1983). The goal in doing so is to identify words that carry the most distinct information; words that appear in nearly every document carry little information (e.g., articles, prepositions). Similarly, words that appear in virtually no documents also carry little information. The most informative word type appears infrequently in the corpus but frequently in a small number of related documents. Common practice is to remove words that carry little information. Higher tf-idf scores reflect this; lower tf-idf scores reflect little discriminatory capacity. Words were retained only if having a tf-idf score of 0.025 or above, slightly above the first quartile.

Documents were then removed if they contained no remaining words (due to only containing low-information words prior to pre-processing), yielding a final 29,631 documents x 142,246 words matrix--a reduction of 354 documents and 47,351 words. This is a smaller reduction that would normally be seen in typical natural language processing applications; however, words were not additionally stemmed and a liberal inclusion strategy was used.

Model parameters

Estimating an LDA model requires specifying the number of topics *ex ante*. As the number of topics increases, the model begins to produce a distinct topic for each event and, as the number increases further, for each document. Selecting the appropriate number of modeled topics is not easily automated. Of course, model selection and whether or not it can be validly automated is also a

³Arabic text is a difficult language to normalize as prefixes, suffixes, and infixes are all used to modify root words. Further, plurals are constructed in multiple ways. Arabic text normalization is an active field of research in natural language processing.

hotly debated area in traditional regression-based research and varies across disciplines. Ultimately, one's substantive modeling goals drive model selection procedures.

Selecting the number of appropriate topics, k , for LDA is a very active field of research. One simple approach is to use the log likelihood for models estimated with a varying number of topics and selecting the maximizing model specification; this approach tends to favor very large numbers of topics. Another approach is to estimate the perplexity, another likelihood-based measure, of a set of held-out test data; the model specification that produces the lowest perplexity is considered to be more likely to have generated the data in the test set. This approach also favors large numbers of topics. A third, more automated approach is hierarchical LDA, an unsupervised method that learns the topical structure of a corpus without specifying the number of topics *a priori* (Blei, Griffiths, Jordan, and Tenenbaum, 2004). Problematically for social scientists, all three of these approaches tend to produce topic models that minimize information retrieval error but frequently produce topics that are less coherent to human interpretation than models in which the number of topics was selected in a stepwise procedure that maximizes interpretation using domain-specific knowledge. For instance, choosing a purely likelihood-based model selection approach frequently produces models with one or two topics or a very large number of topics (Grun and Hornik, 2011).

Similarly, selecting the appropriate values of the two hyperparameters is an ongoing active field of research. Conventional guidelines suggest using a value of $50/k$ for the topic hyperparameter α . Lower values of α increase the probability that documents will be assigned to a single topic; higher values increase the probability that documents will be assigned to multiple topics. Thus, as the number of topics k decreases, the value of α will decrease, which makes logical sense that a greater number of topics increases the likelihood that a topic will exist for each document. The second hyperparameter, η is the topic sparsity parameter and is set to $200/w$ where w is the number of distinct tokens in the modeled corpus. This parameter affects the degree to which words are more likely to be assigned to a few topics or whether the word assignments are likely to be spread more evenly over topics. Given that much of the ongoing research in this area is focused on information retrieval accuracy, conventional values were used.

Data

Data are full-text Arabic news articles drawn from three sources. First, *Al Ahram (The Pyramids)*, the Egyptian state-owned daily newspaper; second, *Al Masry Al Youm (The Egyptian Today)*, a privately owned and independent daily paper, popular with educated Egyptians and Cairenes. The final source is *Al Jazeera*, the pan-Arab satellite network widely viewed as the media source most openly hostile to authoritarian governments (with the exception of the Qatari regime, as the network is based in Doha). Following the ouster of Mubarak and during the parliamentary and presidential elections in Egypt in 2011-2012, *Al Jazeera* has been criticized as openly biased toward the Muslim Brotherhood and the Freedom and Justice Party, the Brotherhood's political party. Although both *Al Ahram* and *Al Masry Al Youm* are daily newspapers, they operate similarly to *Al Jazeera* in updating their websites throughout the day and night as news occurs.

Articles from each source were scraped from their respective websites using software written in the *Scrapy* framework in the Python programming language⁴. For each source, data were collected from the website's archives for the days between 1 November 2010 and 28 February 2011, a 120-day period which encompasses the month before Mohamed Bouazizi's self-immolation, which led to mass protests in Tunisia and Ben Ali's eventual removal, as well as seventeen days following Mubarak's resignation. The beginning date was chosen as *Al Ahram*'s online English archives begin in November 2010 following a redesign that saw much of their archived content moved offline. The ending date was chosen to avoid the revolutionary fervor in much of the Arabic press following Mubarak's resignation unduly influencing the outcome of the topic modeling algorithm discussed below. For all sources, data were collected on local, regional, and international news, opinion, and economic and business news. Entertainment, arts and culture, sports, and other sections were not collected; material in these sections is less likely to be ideologically charged or the subject of major

⁴A web scraper is software that automatically retrieves data from the web by following a set of provided rules and stores it locally. This scraper was written in *Scrapy*, which is a framework written in the Python programming language specifically for writing fast and flexible scrapers. *Scrapy* is available at scrapy.org and source code for this project is available upon request

Table 1: Descriptive statistics, articles per source.

Source	N	Min	Max	Mean	SD
<i>Al Ahrām</i>	10331	7	154	86.1	20
<i>Al Masry Al Youm</i>	11123	0	209	93.5	24
<i>Al Jazeera</i>	8177	33	104	68.1	11.5

narrative contests.

From 28 January until 2 February 2011, Internet access, and to a lesser extent mobile phone access, was effectively shut off across Egypt. Egyptian papers' online editions produced fewer articles during this period. However, reduced media content does not perfectly coincide with the internet blackout for all sources and *Al Jazeera*, whose satellite coverage figured prominently in the uprising, continued to produce content during this time. *Al Masry Al Youm* posted no articles online on 1 February. Topic coverage for this day was interpolated from the preceding and succeeding day using the zoo package for R. The impact of this reduced coverage is starkest for *Al Ahrām's* coverage on 30 January 2011, in which only seven articles were posted online.

Table 1 contains summary statistics for each source. Despite the brief blackout, both *Al Ahrām* and *Al Masry Al Youm* produced the most content in terms of number of articles and in terms of Mean daily output; *Al Jazeera* continued to produce non-reduced content during this period.

Results

Appendix A lists all of the 100 modeled topics with the top ten most likely words for each. Topics were estimated using the lda package (Chang, 2011) for R. The top ten most likely words for each document were translated and the topics were then labeled to provide a short description of topic content. When the association between the words was unclear or seemed conflicting, I read the headlines for the twenty-five most probable articles associated with that topic. Although the vast

majority of topics were easily interpretable, topics 16 and 58 were not.

Results are reported in terms of the proportion of coverage devoted to a given set of topics within sources on a per-day basis. However, this is not equivalent to the proportion of *articles* on a given topic; LDA allows articles to be composed of mixtures of topics; thus, it is possible to model patterns of coverage that would be non-obvious or non-detectable if articles were assigned to a single topic or if words were allowed to be assigned to only one topic. This also allows for the modeling of the interplay between coverage of topics, as discussed below.

Three meta-topics that incorporated the the themes of a) danger and instability, b) regime actions and reforms, and c) opposition and mobilization were created by combining coverage from individual topics related to these themes. The remaining topics are not discussed below. Visualizations for all 100 topics can be found at *URL*.

For each meta-topic, coverage is analyzed for five distinct ``epochs": 1 November 2010 until 13 January 2011, the period from the beginning of data collection until the day before the Tunisian President Ben Ali resigns, a seminal moment in the Arab uprisings. Second, 14 January 2011 until 24 January 2011, the period between Ben Ali's resignation and the first successful Arab revolution until the day before the planned Police Day protests in Egypt that mark the beginning of the 18-day uprising. This period represents the time during which the Mubarak regime would feel threatened by the successful ouster of a fellow autocrat but before mobilization in Egypt has begun. Third, 25 January until 2 February, the first half of the Egyptian uprising. In this period, some of the largest protests of the 18-day uprising occurred, culminating with a ``March of Millions" in Cairo on the first day of February. The next day, on 2 February, communications were restored across the country. Fourth, 3 February until 11 February, the second half of the Egyptian uprising, ending with Mubarak's resignation of the presidency. The decision to restore communications technologies following massive protests in the first half of the uprising should indicate a shift in regime strategy. Finally, 12 February until 28 February, the beginning of the post-Mubarak period and the assumption of power by the Egyptian military. While the periods are uneven in length, they capture ``mini-epochs" in a tumultuous period of Egyptian history and were selected using

substantive knowledge as important and distinct.

Raising the Costs: Danger and Instability

The amount of coverage devoted to danger and instability was operationalized as the proportion of each day's words that were assigned by the algorithm to the following topics: 11 - Bombings & Terrorism, 12 - Health & Hospitals, 13 - Islam, Terrorism, & Religion, 53 - Theft, Crime, & Drugs, 55 - Museums, Theft, & Looting, 81 - Deaths, Woundings, & Accidents, and 86 - Spies & Suicides. Figure 2 displays the overall proportion of coverage to these topics from 1 November 2010 through 28 February 2011. *Al Ahram* is in red, *Al Masry Al Youm* is in green, and *Al Jazeera* is in blue. Each solid vertical line represents a major protest event and the dotted lines represent the beginning, 25 January, and end, 11 February, of the major uprising of 2011⁵.

``Terrorist Organization of 19 Suicide Bombers in Attack on House of Worship ... Interior Minister Reveals in Comprehensive Interview with *Al Ahram* the Details of the Incident in Alexandria"

Front page headline, *Al Ahram*, 25 January 2011⁶.

``Early Rehearsal for 'Day of Rage' .. 12 Demonstrations in Cairo and Governorates to Demand Wage Raises and Bonuses."

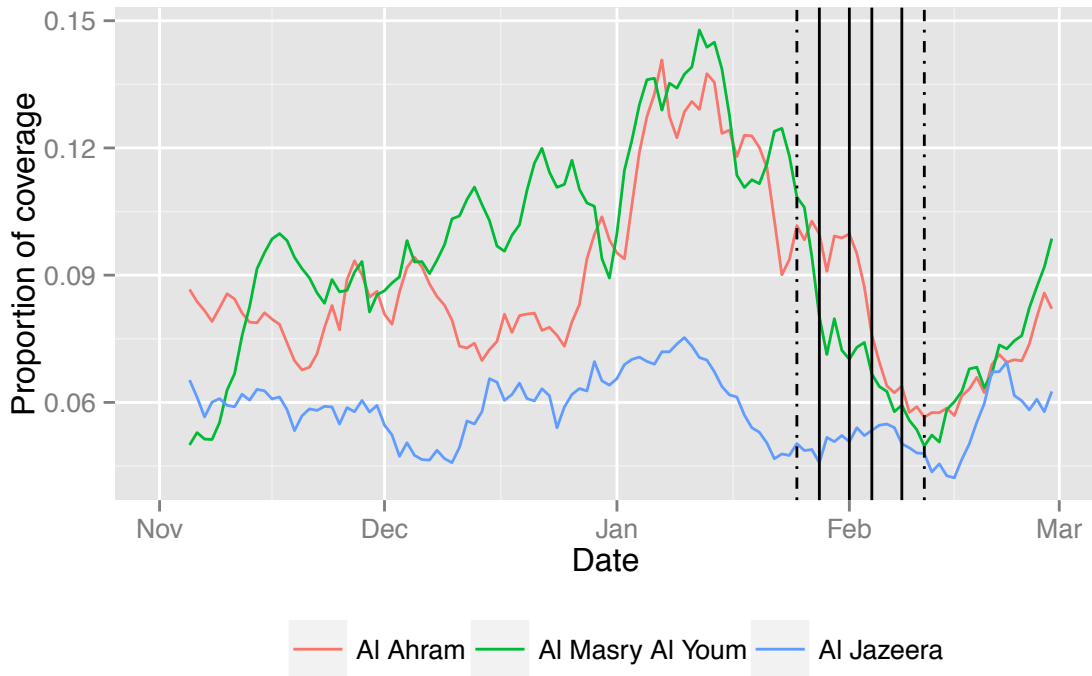
Front page headline, *Al Masry Al Youm*, 25 January 2011⁷.

⁵The question of whether or not the Egyptian uprising constituted a revolution remains an open one; protests against the military continue sporadically. However, it is clear that the 18 days represent a distinct episode of collective action that culminated in regime change.

⁶تنظيم إرهابي من ١٩ انتحاريا لتفجير دور العبادة . . . وزير الداخلية يكشف في حوار شامل ل الأهرام ابعاد حادث الاسكندرية⁶

⁷بروفة مبكرة ل يوم الغضب .. ١٢ مظاهرة في القاهرة والمحافظات للمطالبة بالتعيين ورفع الاجور والمكافآت⁷

Figure 2: Daily proportion of coverage devoted to danger and instability, five-day moving average.



Note: dashed vertical lines indicate the beginning and end of the uprising. Solid vertical lines indicate major protest events.

Figure 2 displays the daily proportion of content devoted to danger and instability for each source using a five-day look-behind simple moving average. It is clear that danger and instability played a major and continuous role in both state and domestic independent media. Large peaks in early January coincide with the bombing of a Coptic church in Alexandria in which 23 people were killed (Reuters, 2011). On average, *Al Masry Al Youm* devotes more of its attention to danger and instability up until the mass uprisings begin, at which point its coverage of these themes drops markedly. Beginning 23 January, two days before protests started, the five-day moving average begins to plummet. At this point, the attention paid to these themes by media outlet switches, and

Table 2: Mean proportion of coverage devoted to danger and instability.

Source	11/1 - 1/13	1/14 - 1/24	1/25 - 2/2	2/3 - 2/11	2/12 - 2/28
<i>Al Ahram</i>	.091	.109	.093	.060	.071
<i>Al Masry Al Youm</i>	.099	.115	.075	.055	.077
<i>Al Jazeera</i>	.061	.052	.051	.051	.058

state media now become the top purveyor of danger and instability themes. *Al Ahram* continues to devote significantly more attention to danger and instability until 1 February, at which point the coverage changes markedly and begins to converge with that of independent and international media.

To test whether the differences observed are statistically significant, the unit of analysis is the topic. There are only 120 observations per topic; subsetting the model further produces the ironic situation of small sample sizes for traditional hypothesis tests. Further, there are no *a priori* reasons to expect that the modeled topics will adhere to distributional assumptions. Therefore, differences in means and confidence intervals for each time period were estimated using a non-parametric bootstrap of 20,000 iterations (Efron and Tibshirani, 1993). All p-values are two-tailed.

Table 2 presents mean coverage for danger and instability for the five epochal time periods. In the period leading up to Ben Ali's resignation of the Tunisian Presidency, *Al Masry Al Youm* devoted more attention to danger and instability than either *Al Jazeera* or *Al Ahram*, ($p < .05$). Following Ben Ali's resignation on 14 January, the difference between *Al Ahram* and *Al Masry Al Youm* is no longer significant, as *Al Ahram* has increased its coverage of these topics and both devote more attention to danger and instability than *Al Jazeera* ($p < .05$).

Once the Egyptian uprising began, *Al Masry Al Youm* continued to devote more attention to danger and instability than *Al Jazeera*, and *Al Ahram* increased its coverage of these topics to exceed that of *Al Masry Al Youm* ($p < .10$). Following the "March of Millions" until the end of the 18-day uprising, the three sources do not significantly differ in their level of attention to danger and

Table 3: Proportion of total coverage of danger and instability on major protest days.

Date	<i>Al Ahram</i>	<i>Al Masry Al Youm</i>	<i>Al Jazeera</i>
25 Jan	.111	.065	.061
28 Jan	.086	.053	.043
31 Jan	.087	.073	.045
1 Feb	.083	(.070)	.036
4 Feb	.097	.070	.063
8 Feb	.057	.061	.045
11 Feb	.042	.045	.037

instability. The period after Mubarak's resignation until 28 February sees a return to pre-uprising patterns, with *Al Ahram* and *Al Masry Al Youm* devoting statistically similar levels of coverage to these topics, both of which exceeded that of *Al Jazeera* ($p < .05$).

Coverage on days of major protest events provides additional confirmation of these patterns; see Table 3 for these figures. On January 24th, the day before the first major protests in Egypt, the state media devotes its second highest amount of attention in a week to danger and instability (13.3%) while independent media devotes the least amount of attention to the issue in almost a month (9.2%, the lowest since 30 December 2010, at 7.4%). The pattern continues on 25 January, the first day of the uprising. State media devote 11.1% of their coverage to instability and danger. Meanwhile independent media's coverage is even lower than the previous day, at 6.5%. This pattern repeats throughout the course of the 18-day uprising. On the date of every major protest event before 8 February (the two Friday protests on 28 January and 4 February as well as 25 January, 31 January, and 1 February), *Al Ahram* devotes more of its coverage to danger and instability than non-state media and *Al Jazeera*, often nearly twice as much.

Beginning on 28 January, the Internet is effectively disconnected for much of Egypt, although the media continue to publish online and in print. Leading up to very large protests on 31 January

and 1 February (the "March of Millions"), state media continue to devote significantly more attention to danger and instability than independent media. One of the largest protests of the uprising occurred on 1 February; this was also the only day in which *Al Masry Al Youm* published no articles on its website. The next day, communications begin to be restored across Egypt and protests continue. Another major protest takes place that Friday, 4 February. Note that the magnitude of effect may be exaggerated for *Al Ahram* on 30 January, as it published only seven articles online that day. Notably, however, those articles heavily favored danger and instability themes.

Although mass protests continue up to and following Mubarak's resignation on 11 February, and despite the fact that violent clashes occurred between anti-regime protesters and pro-Mubarak supporters in this period, no media outlet ever devotes more than 6.6% of its coverage to instability and violence following the penultimate Friday protest on 4 February, which is a turning point in the state media's emphasis of danger and instability. While all media referred to bloody clashes between pro- and anti-regime groups, state media adopted a conciliatory tone, featuring a public apology from regime officials, and publicized concessions to the protesters.

"Shafiq: I Apologize to the Egyptian People for the Events in Tahrir"

Front page headline, *Al Ahram*, 4 February 2011⁸.

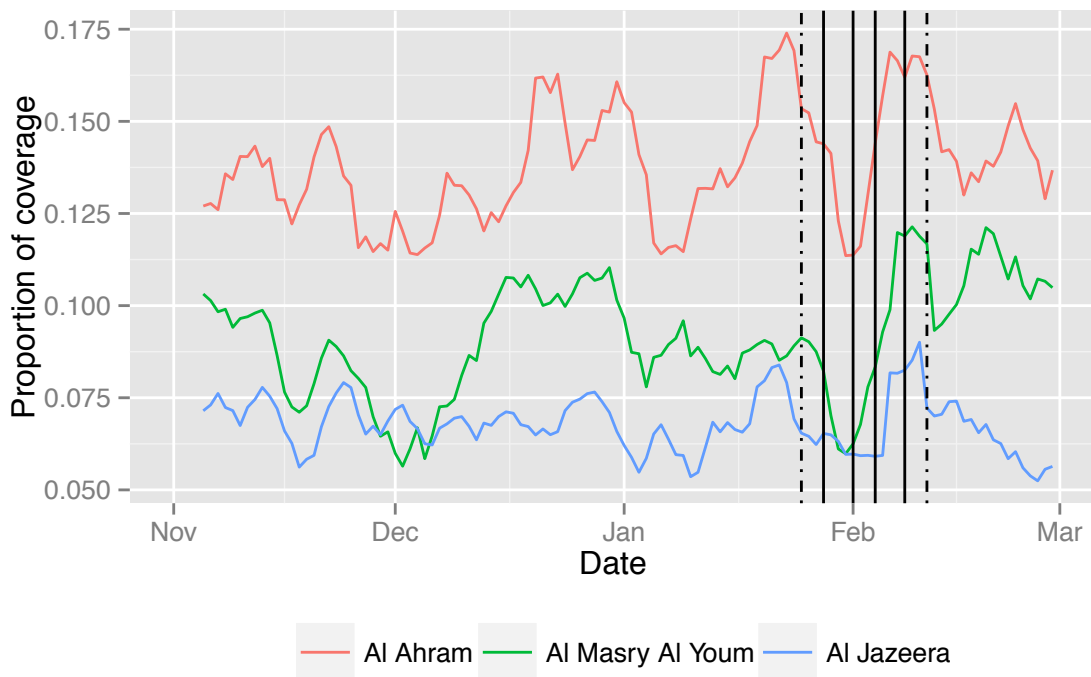
Following this turning point, both *Al Ahram* and *Al Masry Al Youm* approach and sometimes fall beneath *Al Jazeera's* levels of attention to danger and instability; as *Al Jazeera* was widely viewed as a strong supporter of anti-Mubarak protesters, it is unsurprising that *Al Jazeera's* coverage deemphasized dangerous and destabilizing themes. By the eighth of February, three days before Mubarak's eventual resignation, state media are already calling the uprising a revolution, using the protesters' own words.

⁸ شفيق: للشعب المصري عن احداث ميدان التحرير

Lowering the Benefits: Regime actions

The battle for bystanders is not solely about demobilizing the opposition. It also requires assuring supporters that their allegiance will continue to be rewarded. Thus, when faced with the threat of mass collective action, the regime will need to signal that it is cognizant that some changes may be needed but that they will be put in place without the need for mobilization. This appeases bystanders leaning toward participation but who fear instability; it also demonstrates to supporters that the regime continues to be in control.

Figure 3: Daily proportion of coverage devoted to regime actions, five-day moving average.



Note: dashed vertical lines indicate the beginning and end of the uprising. Solid vertical lines indicate major protest events.

The amount of attention paid to regime actions was operationalized as the proportion of each

day's words that were assigned by the algorithm to the following topics: 6 - Business & Entrepreneurship Initiatives, 18 - Central Banking & Finance, 24 - State Land Auctions & Development Projects, 34 - Constitution and Amendments, 42 - Pension Modifications, Tax Exemptions, and Payments, 62 - "Civil Society"⁹, 84 - Mubarak and Official Statements, 91 - Mubarak Meetings and Talks, and 100 - Government Statements about Reform.

Figure 3 presents the overall proportion of coverage to these topics, using a simple five-day moving average (look-behind), from 1 November 2010 through 28 February 2011. *Al Ahrām* is in red, *Al Masry Al Youm* is in green, and *Al Jazeera* is in blue. Each solid vertical line represents a major protest event and the dotted lines represent the beginning, 25 January, and end, 11 February, of the major uprising of 2011. *Al Ahrām*, the state media outlet, devotes more attention across all time points to regime action. Similar to danger and instability themes, state media increases its coverage of regime actions during the interim period between Ben Ali's fall in Tunisia and the beginning of mass protests in Egypt. This is consistent with the proposition that the regime will try to lower the perceived benefits of mobilization. During the same period, independent media maintain a relatively constant level of coverage to the regime's actions.

"Omar Suleiman is Vice President of the Republic and Ahmed Shafiq is Prime Minister"

Front page headline, *Al Ahrām*, 31 January 2011¹⁰.

"The People are Serving the Nation"

Front page headline, *Al Masry Al Youm*, 31 January 2011¹¹.

Coverage patterns of regime actions diverge sharply during 18-day uprising, with state media devoting substantially more attention to regime actions than domestic independent or international

⁹Nearly all of the most likely articles associated with this topic were from *Al Ahrām* and represented a common feature of modern authoritarianism -- co-opted civil society groups, dubbed governmental non-governmental organizations.

¹⁰ عمر سليمان نائباً لرئيس الجمهورية و احمد شفيق رئيساً للوزراء

¹¹ الشعب في خدمة الوطن

Table 4: Mean proportion of coverage devoted to regime actions.

Source	11/1 - 1/13	1/14 - 1/24	1/25 - 2/2	2/3 - 2/11	2/12 - 2/28
<i>Al Ahrām</i>	.133	.157	.129	.165	.138
<i>Al Masry Al Youm</i>	.089	.089	.071	.109	.109
<i>Al Jazeera</i>	.068	.072	.063	.077	.061

media. The size of the *Al Ahrām* effect for 30 January is almost certainly exaggerated due to limited media output that day. However, the divergence is very clear both visually, in Figure 3 as well as in Table 4. It is not until the midpoint of the uprising, following the "March of Millions" on 1 February that *Al Masry Al Youm* and *Al Jazeera* begin to increase their coverage of the regime's actions again. During this time Gamal Mubarak, the president's son, and the rest of the leadership of Mubarak's National Democratic Party (NDP) resigned their posts and constitutional amendments were proposed.

Despite the fact that the protesters directed their actions firmly at Mubarak's government and chanted for the downfall of the regime, it is clear that non-state media did not focus heavily on the actions of the regime itself. While the Mubarak regime had a long history of adeptly proposing just enough reforms to quiet a dissatisfied polity, it appears that it failed in spreading this message to non-state media at the moment at which it needed it most. During the first half of the uprising, *Al Masry Al Youm* and *Al Jazeera* devote much less coverage to the regime's actions, leaving *Al Ahrām* as the outlier once again in selling an unconvincing narrative to bystanders. Following Mubarak's resignation on 11 February, state and non-state media converge once again to pre-uprising levels and difference in levels, while *Al Jazeera's* coverage drops off once again.

"Sharif, Gamal Mubarak, Ezma, Ezz, Shahab, and Hilal Resign From NDP"

Front page headline, *Al Ahrām*, 6 February 2011¹².

¹² استقالة الشريف وجمال مبارك وعزمي وعز وشهاب وهلال من امانة الوطني

Table 5: Proportion of total coverage of regime actions on major protest days.

Date	<i>Al Ahram</i>	<i>Al Masry Al Youm</i>	<i>Al Jazeera</i>
25 Jan	.143	.100	.057
28 Jan	.157	.070	.076
31 Jan	.105	.060	.055
1 Feb	.132	(78)	.062
4 Feb	.121	.082	.045
8 Feb	.171	.095	.062
11 Feb	.167	.098	.085

“Youth of 25 January Topples Gamal Mubarak”

Front page headline, *Al Masry Al Youm*, 31 January 2011¹³.

Differences in means and confidence intervals were constructed using a non-parametric bootstrap of 20,000 iterations to confirm the patterns discussed above. In both periods leading up to the uprising's beginning, *Al Ahram* devoted significantly more coverage to regime action than *Al Masry Al Youm* ($p < .05$), which in turn devoted more coverage than *Al Jazeera* ($p < .05$). From 25 January until 2 February, *Al Masry Al Youm* and *Al Jazeera* do not differ in their coverage of regime action, but *Al Ahram* continues to devote significantly more attention to the topic ($p < .05$). The original ordering of $Al Ahram > Al Masry Al Youm > Al Jazeera$ returns following Mubarak's resignation ($p < .05$). Coverage on days of major protest events, seen in Table 5, confirm this finding.

¹³ شباب ٢٥ يناير يطيح جمال مبارك

Invalidation: Opposition and Mobilization

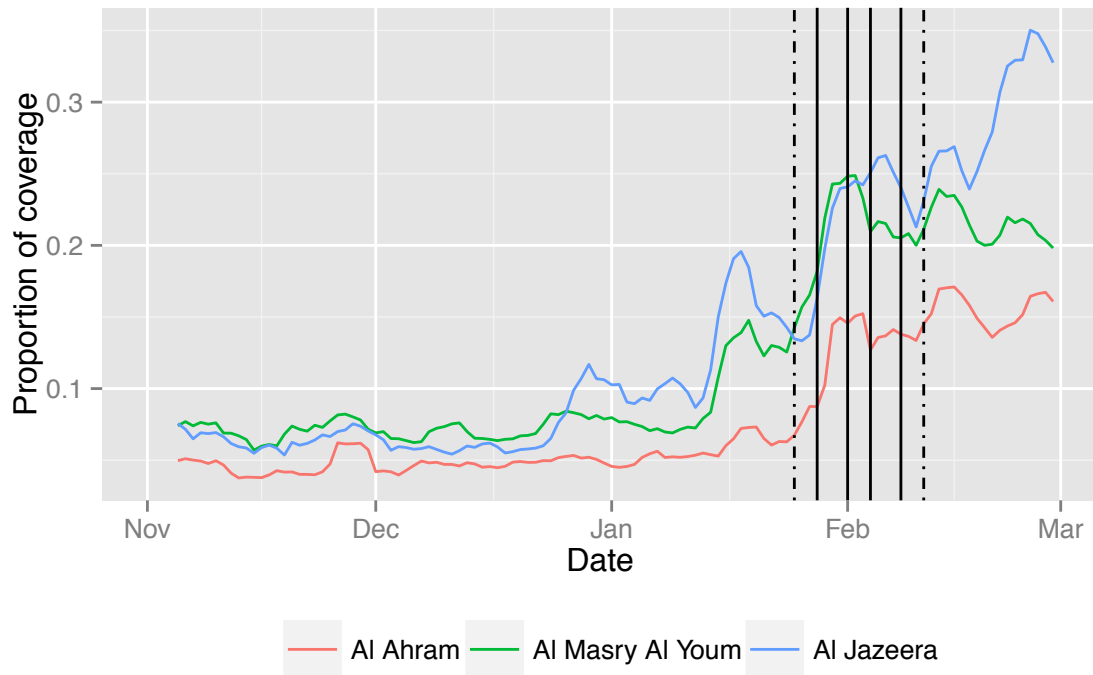
The third component of the battle for bystanders is the coverage of mobilization itself. The regime would prefer to invalidate perceptions of political opportunities by downplaying instances of successful mobilization to avoid demonstration effects. Suppressing common knowledge about successful mobilization is the most effective way to decrease future mobilization and is less costly than trying to convince bystanders that ongoing collective action is dangerous, destabilizing, or unnecessary. This includes coverage of domestic mobilization, opposition activities, and, as protests were ongoing around the Arab world, successful mobilization in neighboring states.

The amount of attention paid to regime actions was operationalized as the proportion of each day's words that were assigned by the algorithm to the following topics: 1: Revolution & the Army, 9 - Libya, 30 - Yemen, 37 - Demonstrations, 65 - Algerian Demonstrations, 67 - Labor Strikes, Work Stoppages, & Unions, 73 - Khaled Saeed Movement, 74 - Muslim Brotherhood, Baradei, and Opposition, 93 - Tunisia, 96 - Lawyers Syndicate & Unions. Figure 4 presents the overall proportion of coverage to these topics from 1 November 2010 through 28 February 2011. *Al Ahram* is in red, *Al Masry Al Youm* is in green, and *Al Jazeera* is in blue. Each solid vertical line represents a major protest event and the dotted lines represent the beginning, 25 January, and end, 11 February, of the major uprising of 2011.

Before protests begin in Egypt, *Al Masry Al Youm*, the non-state domestic media outlet, devotes more coverage to opposition actors and mobilization than state media, though not staggeringly so, as seen in Figure 4. However, as with danger and instability and regime actions, the battle for bystanders begins in mid January and continues apace until Mubarak's resignation on 11 February. Beginning with Ben Ali's resignation on January 14, *Al Masry Al Youm* and *Al Jazeera* expand their coverage of opposition and mobilization by almost double. It is striking how little coverage *Al Ahram*, the state media outlet, devotes to these themes during Tunisia's revolution--a clear indicator that the Mubarak regime felt threatened by the possibility of the diffusion of mobilization.

Yet, the coverage of *Al Masry Al Youm* and *Al Jazeera* skyrocketed to more than 25% of all content on some days. Again, some of the effects may be exaggerated between 28 January and

Figure 4: Daily proportion of coverage devoted to opposition and mobilization, five-day moving average.



Note: dashed vertical lines indicate the beginning and end of the uprising. Solid vertical lines indicate major protest events.

1 February due to limited internet publishing, however, they are consistent with the effects seen before and after this period. Omitting 30 January as a possible outlier due to small sample size, state media never devote more than 16.5% of all coverage to opposition and mobilization themes during the entire 18-day uprising. This is remarkable when contrasted with *Al Masry Al Youm* and *Al Jazeera*, which never devote *less* than 15.4% of their coverage to these themes (and Mean in the 20s).

To confirm the above findings, differences in means (displayed in Table 6) and confidence

Table 6: Mean proportion of coverage devoted to opposition and mobilization.

Source	11/1 - 1/13	1/14 - 1/24	1/25 - 2/2	2/3 - 2/11	2/12 - 2/28
<i>Al Ahrām</i>	.048	.066	126	144	157
<i>Al Masry Al Youm</i>	.072	133	226	215	212
<i>Al Jazeera</i>	.074	166	216	243	302

intervals were constructed using a non-parametric bootstrap of 20,000 iterations. During the first period, prior to Ben Ali's resignation, *Al Masry Al Youm* and *Al Jazeera* cover opposition and participation themes significantly more than *Al Ahrām* ($p < .05$) but do not significantly differ from each other. In the interim period before 25 January but after Ben Ali's resignation, *Al Jazeera* is already focusing more than either source ($p < .05$) and *Al Masry Al Youm* is covering opposition and participation themes significantly more than *Al Ahrām* ($p < .05$). During the first half of the 18-day uprising in Egypt, *Al Jazeera* and *Al Masry Al Youm* do not differ in their coverage, but both devote significantly more attention to these themes than *Al Ahrām* ($p < .05$). The same pattern holds for remainder of the uprising and after, although the gap between *Al Ahrām* and *Al Masry Al Youm* gets smaller over time: the 95% confidence intervals of the difference in means across the final three epochal periods are 0.05 - 0.15, 0.05 - 0.09, and 0.04 - 0.07, respectively. Following Mubarak's resignation, protests continued across the Middle East and North Africa and *Al Jazeera* continued with the story, while domestic media remained domestically focused. Interestingly, by 8 February, three days before Mubarak's eventual resignation, state media were already calling the events a "revolution."

"25 January Revolution Attempts to Form a Coalition"

Front page headline, *Al Ahrām*, 8 February 2011¹⁴.

¹⁴ محاولات لتشكيل ائتلاف ثورة ٢٥ يناير

Table 7: Proportion of total coverage of opposition and mobilization on major protest days.

Date	<i>Al Ahram</i>	<i>Al Masry Al Youm</i>	<i>Al Jazeera</i>
25 Jan	.077	.209	.102
28 Jan	.091	.199	.257
31 Jan	.121	.204	.221
1 Feb	.097	(.203)	.213
4 Feb	.162	.210	.286
8 Feb	.121	.227	.201
11 Feb	.161	.252	.312

The Mubarak regime seemed utterly unable to contain coverage of the opposition or mobilization right from the very beginning of the Arab uprisings. By attempting to minimize coverage of successful mobilization while emphasizing its own reform actions and the danger of mobilizing, the regime was already losing the battle of narratives by the time the Egyptian uprising began in earnest on 25 January. The sharp divergence seen between independent media, both domestic and international, and state media in content of coverage is a signal that the battle for bystanders is underway and signals the opening of a political opportunity for mobilization. Despite recent elections widely regarded as incredibly fraudulent and widely held grievances about an under-performing economy, signals of opening political opportunities did not appear in the press until mid-January 2011, as will be discussed below.

Narrative Interplay

Taken in isolation, findings for each of the three themes provide support that a narrative battle was occurring in the Egyptian media prior to and during the January-February mass uprising, providing signals to non-committed bystanders that the regime's control of the interpretation of political reality

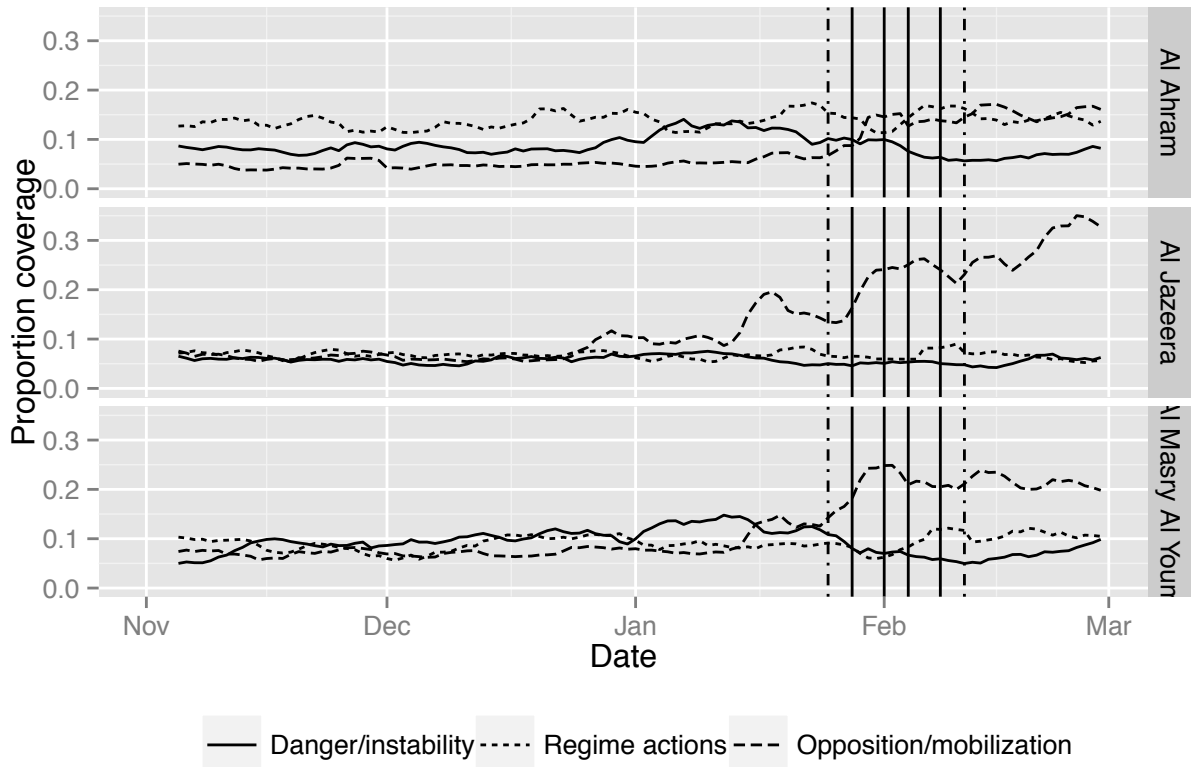
was in serious jeopardy and that political opportunities for collective action were opening. The interplay between these themes both within and between media sources is an additional important source of information for bystanders. Figure 5 presents the entire timeline for each of the three sources using a five-day simple moving average (look-behind). The solid line in each plot is the proportion of coverage for that source devoted to danger and instability themes, the dotted line in each plot is the proportion of coverage devoted to regime actions, and the dashed line in each plot is the proportion of coverage devoted to opposition and mobilization. The top plot represents state media, the middle plot is international media, and the bottom plot is independent media.

It is clear that independent domestic and international media led the way in establishing coverage that focuses on mobilization and opposition, first and foremost, followed by regime actions, and then by danger and instability. This pattern is established in *Al Jazeera* potentially starting from early January, but in earnest from mid-January onwards. As the uprising in Egypt progresses, the difference in the amount of coverage is stark. *Al Masry Al Youm* follows a similar, if slightly delayed pattern, although mobilization and opposition are established as the dominant topics before the uprising begins *en masse* on 25 January. The only date which mobilization and opposition are not the most prominent topics is on 7 February, when regime actions briefly take the focus following mass resignations in the NDP.

In contrast, the *Al Ahrām* data present a picture of a regime struggling to juggle multiple narratives and failing to win control of the narrative as mobilizations continued across the country. *Al Ahrām* does not begin to resemble the other media outlets until after 4 February, when it the danger and instability narrative is largely abandoned, more than a week after *Al Masry Al Youm* and nearly two weeks *Al Jazeera* had already shifted their focus to successful mobilization.

Why should we believe that these results provide support for an argument that battles over narratives in the media send signals to non-committed bystanders and influence collective action decision-making? Given that the data under consideration are newspaper articles, it is possible that the revealed patterns are merely representative of the underlying data-generating process: current events are newsworthy and covered by the media. This is certainly at least partially the case.

Figure 5: Proportion of daily coverage to three narratives.



Note: dashed vertical lines indicate the beginning and end of the uprising. Solid vertical lines indicate major protest events.

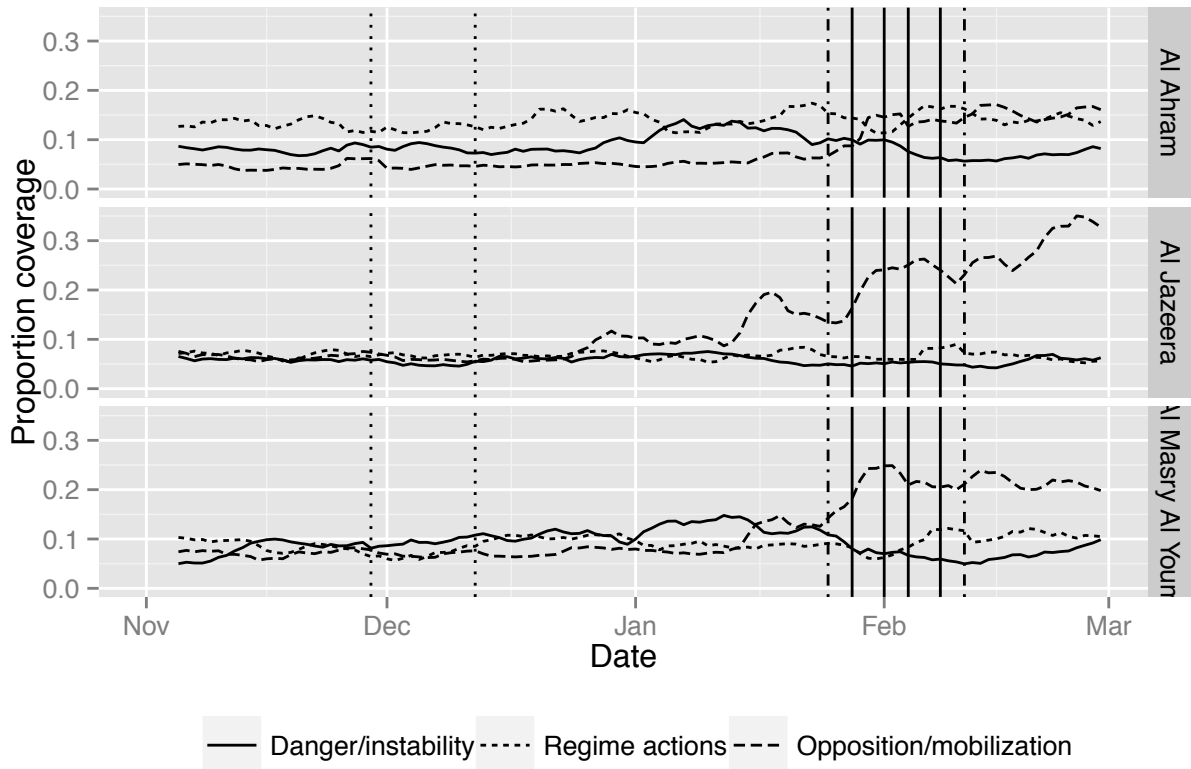
However, there are two reasons to have confidence in the argument presented above: a) timing and b) prior protests.

Timing. As argued above, both *Al Masry Al Youm* and *Al Jazeera* presented a narrative of events in Egypt that focused on successful mobilization and opposition far more than those of danger, instability, or attempts by the Mubarak regime to appease protesters. These are the dominant narrative patterns prior to the beginning of the uprising in Egypt and remain so through its end. *Al Ahrām*, on the other hand, eventually adopts the same narrative stance, but only after more than a week of protests featuring thousands and *after* the regime's last-ditch effort to disrupt communication through cellular and internet disconnections. It seems that the first few days of February, after the "March of Millions" were the point at which the regime abandoned its attempt to control the narrative and frighten protesters.

The fact that both non-state and international media narratives remain relatively constant while state media demonstrates movement toward their positions supports the hypothesis that a contest was taking place--a contest that the regime visibly lost. Not only was state media unable to portray mobilization as too costly to prevent protests, it was also unable to control the dominant interpretation of events. If the regime was able to exert significant control over the media or if its interpretation of events were favored, one would expect that *Al Ahrām* would lead the way in establishing eventual coverage patterns; this is the opposite of what happened.

Prior protests. Figure 6 displays the same data as Figure 5, with two additions. The vertical black bars represent two protest events in Egypt that received international press coverage prior to the start of the Arab uprisings. The first, on 29 November 2010, involved multiple protests across Egypt over election fraud and widespread losses by the Muslim Brotherhood. At least two protesters died in riots and multiple fires were set (Associated Press, 2010; BBC, 2010). The second, on 12 December 2010, were also related to election results and took place mainly in Cairo and involved hundreds (Mikhail, 2010). Despite the protests being large enough to receive international attention from the Associated Press, Reuters, and BBC, they did not spark large-scale collective action like that seen less than a month later.

Figure 6: Proportion of daily coverage to three narratives, with prior protests.



Note: dashed vertical lines indicate the beginning and end of the uprising. Solid vertical lines indicate major protest events. Dotted vertical lines indicate minor protest events.

Media coverage patterns do not exhibit signs of a battle over the narrative, either. Using non-parametric bootstrapped confidence intervals of 20,000 iterations to test for differences in coverage for the week preceding the November protests, *Al Ahram* and *Al Masry Al Youm* do not significantly differ in their coverage of danger and instability, although both cover these themes significantly more than *Al Jazeera* ($p < .05$). As is normal, *Al Ahram* devotes more coverage to regime actions than either *Al Masry Al Youm* or *Al Jazeera* ($p < .05$). *Al Masry Al Youm* and *Al Jazeera* devote slightly more coverage to opposition and mobilization ($p < .05$) than *Al Ahram*, but not

substantially so (95% CI: .003 - .03).

The December protests follow a similar pattern; *Al Masry Al Youm*, domestic independent media, actually devote more coverage to danger and instability than either state media or international media ($p < .05$). State media devote more coverage to regime actions than other sources ($p < .05$), but the magnitude of difference is stable (95% CI: .04 - .06). *Al Masry Al Youm* devote significantly more coverage to opposition and mobilization than state media ($p < .05$) and *Al Jazeera* devote more coverage than either of the other two sources ($p < .05$), though the differences are in the range of 1-3 percent.

Visually, it is clear that the large swings in narrative did not occur surrounding either of these protest events. The regime did not disconnect the internet or mobile service. No battle for bystanders was underway and the interpretation of the events was not widely contested. In other words, the most fraudulent elections in years were seen as business as usual and grievances arising from them were not perceived to be a serious threat by the government or by its critics and opponents.

Conclusion

The media do not cause revolutions. As a voluminous literature on social movements, revolutions, and contentious politics makes clear, resources and social networks matter. Recruiting participants, via personal interaction, online social networks, religious groups, or other avenues is how large-scale movements are made. However, spontaneous participation in mass protests is a major feature of large-scale collective action. Individuals seek some way of assessing the risks and benefits of their participation. In information-poor environments like those cultivated by autocratic regimes, individuals need some way of assessing the distribution of political preferences as well as the likelihood of success of pro- and anti-regime action.

The mass media provide a low-cost way of obtaining this information and help cultivate common knowledge. The regime's actions with respect to the interpretation of events in the media provide signals about the opening or closing of political opportunities. The interplay of narratives in state and non-state media outlets provide information about how well the regime is able to control the

flow and presentation of information.

I have shown that the Mubarak regime in Egypt attempted to get in front of potential anti-government collective action by emphasizing themes of danger and instability while also focusing heavily on its own "reform" activities. In doing so, it attempted to simultaneously raise the perceived costs of mobilization while lowering its benefits. Caught in a struggle between appeasing its supporters and demobilizing its allies, the regime was unable to convincingly accomplish both as opposition media, both foreign and domestic, focused on successful opposition and mobilization, while deemphasizing danger and instability and virtually ignoring the actions of the Mubarak regime. It seems that the narrative media battle for non-committed bystanders was lost by the government before the first protests in Egypt on Police Day in January 2011.

Although much discussion has been generated about the role of social media such as Facebook and Twitter in the Arab uprisings (Farrell, 2012; Howard, Duffy, Freelon, Hussain, Mari, and Mazaid, 2011; Lotan, Graeff, Ananny, Gaffney, Pearce, and boyd, 2011; Lynch, 2011; Papacharissi and de Fatima Oliveira, 2012), I have not included social media in my analysis. First, it is beyond the ambition of this paper to offer a complete theory of collective action or even a complete explanation of the 18-day uprising in Egypt. Second, preliminary (though non-representative) survey data collected from protesters in Cairo's Tahrir Square indicate that protesters were almost one and a half times as likely to regularly use print media than Facebook (63% to 43%, Wilson and Dunn (2011). Thus, while social media's role in the uprising is still a matter of debate, it is not the case that social media were the primary source of news for many Egyptians. Brym, Godbout, Hoffbauer, Menard, and Zhang (2012) also find, using survey data collected by Gallup in March/April 2011, that protesters were more likely to consume *both* state and independent news than non-protesters -- and that traditional media sources were more important sources of news about the uprising than new media.

Finally, I have demonstrated that although the mass media in authoritarian societies may be disregarded by some as merely propaganda, this is not the case. The mass media in modern autocracies do not have the luxury of publishing news that is completely at odds with reality; they operate

alongside independent media, international media that broadcast over satellite, and the internet. How agile regimes are in managing not only information but also its dominant interpretation will become increasingly important as informational borders become more porous and may well help decide the fate of many regimes in the near future.

References

- Amenta, Edwin, Neal Caren, Sheera Joy Olasky, and James E. Stobaugh. 2009. "All the Movements Fit to Print: Who, What, When, Where, and Why SMO Families Appeared in the New York Times in the Twentieth Century." *American Sociological Review* 74:636--656.
- Andrews, Kenneth T. and Michael Biggs. 2006. "The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960 Sit-Ins." *American Sociological Review* 71:752--777.
- Andrews, K. T. and N. Caren. 2010. "Making the News: Movement Organizations, Media Attention, and the Public Agenda." *American Sociological Review* 75:841--866.
- Associated Press. 2010. "Egypt: Protesters Accuse Ruling Party of Voter Fraud." *The New York Times* .
- Baum, Matthew A. and Tim Groeling. 2008. "New Media and the Polarization of American Political Discourse." *Political Communication* 25:345--365.
- BBC. 2010. "Egypt Brothers 'lose most seats'." *BBC* .
- Benford, Robert D. and David A. Snow. 2000. "Framing Processes and Social Movements: An Overview and Assessment." *Annual Review of Sociology* 26:611--639.
- Blaydes, Lisa. 2006. "Who Votes in Authoritarian Elections and Why? Determinants of Voter Turnout in Contemporary Egypt." Philadelphia, PA.
- Blei, David M. 2011. "Introduction to Probabilistic Topic Models."
- Blei, David M, Thomas L Griffiths, Michael I Jordan, and Joshua B Tenenbaum. 2004. "Hierarchical topic models and the nested Chinese restaurant process." *Advances in Neural Information Processing Systems* 16:2003.

- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent dirichlet allocation." *Journal of Machine Learning Research* 3:993–1022.
- Brosius, Hans-Bernd and Gabriel Weimann. 1996. "Who Sets the Agenda." *Communication Research* 23:561--580.
- Brownlee, Jason. 2007. *Authoritarianism in an Age of Democratization*. Cambridge: Cambridge University Press.
- Brym, Robert, Melissa Godbout, Andreas Hoffbauer, Gabe Menard, and Tony Huiquan Zhang. 2012. "Demonstrators, Sympathetic Onlookers, and Social Media in the Arab Spring: The 2011 Egyptian Uprising."
- Chang, Jonathan. 2011. "lda: Collapsed Gibbs sampling methods for topic models."
- Chang, Jonathan, Jordan Boyd-Graber, Sean M. Gerrish, Chong Wang, and David M Blei. 2009. "Reading Tea Leaves : How Humans Interpret Topic Models." *Neural Information Processing Systems* 31:1--9.
- Chwe, Michael Suk-Young. 1999. "Structure and Strategy in Collective Action." *American Journal of Sociology* 105:128--156. ArticleType: research-article / Full publication date: July 1999 / Copyright 1999 The University of Chicago Press.
- Coyne, Christopher J. and Peter T. Leeson. 2009. "Media as a Mechanism of Institutional Change and Reinforcement." *KYKLOS* 62:1--14.
- Efron, Brad and Robert J. Tibshirani. 1993. *An introduction to the bootstrap*. London: Chapman and Hall.
- Farrell, Henry. 2012. "The Internet's Consequences for Politics." *Annual Review of Political Science* 15:15--52.

- Gamson, William A. 2007. "Bystanders, Public Opinion, and the Media." In *The Blackwell Companion to Social Movements*, edited by David A. Snow, Sarah A. Soule, and Hanspeter Kriesi, p. 242–261. Blackwell Publishing Ltd.
- Gamson, William A., David Croteau, William Hoynes, and Theodore Sasson. 1992. "Media Images and the Social Construction of Reality." *Annual Review of Sociology* 18:373--393.
- Gamson, William A. and David S. Meyer. 1996. "Framing political opportunity." In *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*, edited by Doug McAdam, John D. McCarthy, and Mayer N. Zald, pp. 275--290. Cambridge: Cambridge University Press.
- Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83:1420--1443.
- Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18:1--35.
- Grun, Bettina and Kurt Hornik. 2011. "topicmodels: An R Package for Fitting Topic Models." *Journal of Statistical Software* 40.
- Hall, David, Daniel Jurafsky, and Christopher D Manning. 2008. "Studying the history of ideas using topic models." pp. 363--371. Association for Computational Linguistics.
- Hopkins, Daniel J and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54:229--247.
- Howard, Philip N., Aiden Duffy, Deen Freelon, Muzammil M. Hussain, Will Mari, and Marwa Mazaid. 2011. "Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring?" Technical Report 2011.1, Project on Information Technology & Political Islam.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *The Journal of Economic Perspectives* 5:193--206.

ArticleType: research-article / Full publication date: Winter, 1991 / Copyright 1991 American Economic Association.

Koopmans, Ruud and Susan Olzak. 2004. "Discursive Opportunities and the Evolution of Right-Wing Violence in Germany." *American Journal of Sociology* 110:198--230.

Kuran, Timur. 1991. "Now out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44:7--48.

Kuran, Timur. 1997. *Private truths, public lies: the social consequences of preference falsification*. Harvard University Press.

Kurzban, Charles. 1996. "Structural Opportunity and Perceived Opportunity in Social-Movement Theory: The Iranian Revolution of 1979." *American Sociological Review* 61:153--170.

Larcinese, V., R. Puglisi, and J.M. Snyder Jr. 2011. "Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers." *Journal of Public Economics* .

Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge University Press, 1 edition.

Lichbach, Mark I. 1994. "Rethinking Rationality and Rebellion Theories of Collective Action and Problems of Collective Dissent." *Rationality and Society* 6:8--39.

Lichbach, Mark Irving. 1998. *The Rebel's Dilemma*. University of Michigan Press.

Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91." *World Politics* 47:42--101.

Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and danah boyd. 2011. "The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions." *International Journal of Communication* 5:1375--1405.

- Lynch, Marc. 2011. "After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State." *Perspectives on Politics* 9:301--310.
- Macy, Michael W. 1991. "Chains of Cooperation: Threshold Effects in Collective Action." *American Sociological Review* 56:730--747.
- Maher, Thomas V. 2010. "Threat, Resistance, and Collective Action: The Cases of Sobibór, Treblinka, and Auschwitz." *American Sociological Review* 75:252--272.
- McCombs, Maxwell E. and Donald L. Shaw. 1972. "The Agenda-Setting Function of Mass Media." *Public Opinion Quarterly* 36:176 --187.
- McCombs, Maxwell E and Donald L Shaw. 1993. "The Evolution of Agenda-Setting Research: Twenty-Five Years in the Marketplace of Ideas." *Journal of Communication* 43:58--67.
- Meyer, David S. 2004. "Protest and Political Opportunities." *Annual Review of Sociology* 30:125-145.
- Mikhail, Sarah. 2010. "Egyptians protest parliamentary vote results." *Reuters* .
- Newman, David J. 2006. "Probabilistic topic decomposition of an eighteenth-century american newspaper." *Journal of the American Society Information Science Technology* 57:2006.
- Oliver, Pamela E. and Gregory M. Maney. 2000. "Political Processes and Local Newspaper Coverage of Protest Events: From Selection Bias to Triadic Interactions1." *American Journal of Sociology* 106:463--505.
- Oliver, Pamela E and Gerald Marwell. 2001. "Whatever Happened to Critical Mass Theory? A Retrospective and Assessment." *Sociological Theory* 19:292--311.
- Oliver, Pamela E. and Daniel J. Meyer. 1999. "How Events Enter the Public Sphere: Conflict, Location, and Sponsorship in Local Newspaper Coverage of Public Events." *American Journal of Sociology* 105:38--87.

- Papacharissi, Zizi and Maria de Fatima Oliveira. 2012. "Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt." *Journal of Communication* .
- Pfaff, Steven. 2006. *Exit-voice dynamics and the collapse of East Germany: the crisis of Leninism and the revolution of 1989*. Duke University Press.
- Polletta, Francesca. 2008. "Frames and Their Consequences." In *The Oxford Handbook of Contextual Political Analysis*, edited by M. Kai Ho, Robert Goodin, and Charles Tilly, volume 1, pp. 187--210. Oxford, England: Oxford University Press, 1 edition.
- Reuters. 2011. "Egypt church blast death toll rises to 23." *Reuters* .
- Roscigno, Vincent J. and William F. Danaher. 2001. "Media and Mobilization: The Case of Radio and Southern Textile Worker Insurgency, 1929 to 1934." *American Sociological Review* 66:21.
- Salton, Gerard, Edward A. Fox, and Harry Wu. 1983. "Extended Boolean information retrieval." *Commun. ACM* 26:1022--1036.
- Sawyers, Traci M. and David S. Meyer. 1999. "Missed Opportunities: Social Movement Abeyance and Public Policy." *Social Problems* 46:187--206.
- Scheufele, D. A. 1999. "Framing as a theory of media effects." *Journal of Communication* 49:103-122.
- Schrodt, Philip A. 2010. "A Brief Introduction to Text as Data."
- Schudson, Michael. 2002. "The News Media as Political Institutions." *Annual Review of Political Science* 5:249--269.
- Slater, Dan. 2009. "Revolutions, Crackdowns, and Quiescence: Communal Elites and Democratic Mobilization in Southeast Asia." *American Journal of Sociology* 115:203--254.

- Snow, David A. 2007. "Framing Processes, Ideology, and Discursive Fields." In *The Blackwell Companion to Social Movements*, edited by David A. Snow, Sarah A. Soule, and Hanspeter Kriesi, p. 380–412. Blackwell Publishing Ltd.
- Steyvers, Mark and Tom Griffiths. 2007. "Probabilistic Topic Models." In *Latent Semantic Analysis: A Road to Meaning*, edited by T. Landauer, D. McNamara, S. Dennis, and W. Kintsch.
- Tsfati, Yariv. 2003. "Media Skepticism and Climate of Opinion Perception." *International Journal of Public Opinion Research* 15:65--82.
- Wilson, Christopher and Alexandra Dunn. 2011. "Digital Media in the Egyptian Revolution: Descriptive Analysis from the Tahrir Data Set." *International Journal of Communication* 5:1248-1272.
- Zhao, Yuezhi. 1998. *Media, Market, and Democracy in China: Between the Party Line and the Bottom Line*. University of Illinois Press.

Appendix: Modeled Topics and Ten Most Likely Words

Topic Label	Ten Most Likely Words
1 Revolution & Army	ثورة مبارك جيش مسلحة ميدان شهداء تنحي انتقالية ثوار متظاهرين revolution Mubarak army armed square martyrs (step down) transition rebels demonstrators
2 Omar Effendi & Orascom	شركة شركات جني بيع هيئة صفقة قابضة حديد شراء لشركة company companies reap sale body/committee deal holding iron buy (for company)
3 Elections & Fraud	حزب انتخابات وفد نواب تزوير جولة انتخابية انسحاب مقاعد مؤتمر party election Wafd Congress fraud election round withdrawal seat meeting/conference
4 Saudis & Gulf	سعودية قطر ملك سعودي امير شيخ امارات خليج مملكة حمد Saudi Qatar king Saudi prince sheikh Emirates Gulf kingdom Hamad
5 Afghanistan & Pakistan	افغانستان اميركية قاعدة باكستان طالبان حلف ناتو افغانية ارباب شمال Afghanistan America al-Qaeda Pakistan Taliban COALITION NATO Afghan terrorism Afghani
6 Business & Entrepreneurship Initiatives	مشروعات قطاع استثمار تطوير شركات استثمارات مجالات زيادة خدمات صناعة projects sector investment development companies investments areas/sectors growth services finance
7 Problems in Education	تعليم مدارس مدرسة تربية تعليمية طلاب اطفال جامعي تطوير ثانوية education teachers school education/upbringing educational students children college secondary development
8 Art	ثورة سادات مسرح فنان رسم فضاء مثقف ابداع لوحات لون revolution Sadat theater/scene artist draw space cultured images creativity color
9 Libya	قذافي ليبيا ليبيا ليبيا ليبيا طرابلس معمر بنغازي متظاهرين ليبيا عقيد Gaddafi Libya Libyan Libya Tripoli Muammar Benghazi demonstrators Libyan colonel
10 Football/Soccer	مركز دقيقة لقب نقطة برازيل اسبانية صحراء اسبانيا اسباني مباريات center/site minute title point Brazil Spain Spanish Sahara Spanish matches

Topic Label	Ten Most Likely Words
11 Bombings & Terrorism	جيش هجوم شرطة قتلي انفجار شخصا هجمات جنود شمال جنوب army attack police killed explosion attacks people soldiers north killed
12 Health & Hospitals	صحة علاج مستشفى طبية مستشفيات اطباء مرض حالات مرضي امراض health treatment hospital medical hospitals doctors illness cases ill disease
13 Islam, Terrorism, & Religion	اسلام ارهاب جريمة فتنة مسجد مساجد اديان ارهابية مسيحية قران Islam terrorism trial/disorder crime mosque mosques religions terrorist Christians <i>Qur'an</i>
14 Pleading Op-Eds	وطن انت عبدالناصر انتم وسوف فلماذا ملك جريدة وانت اكتب nation/national you Nasser you (will be) why king newspaper (and you) wrote
15 Natural Resources	غاز نפט بترول انتاج طاقة برمبل شركة يوميا خام وقود gas oil petrol production energy barrel company daily ore fuel
16 <i>Uninterpretable</i>	جوهر حراك عجز خريطة تحرر خصوصية دوائر بيمة نقطة تقسيم essence/core movement inability/paralysis map/plan liberation/Tahrir privacy/secretcy departments environment point/period division/separation
17 Family & Women	اسرة منزل زواج زوجت سيدة زوجة عائلة طفل زوجها اسر family home marriage married woman wife family child (her husband) family
18 Central Banking & Finance	بنك بنوك مركزي دولار جني اموال صندوق صرف تمويل نقد bank banks centers dollar earn money fund exchange finance criticism/money/cash
19 <i>Uninterpretable</i>	ولد تضامن خبز توزيع تموين دقيق عزيز بوتاجاز اسرة سوداء born/son solidarity bread/sow distribute provision accurate/careful Aziz/precious cooker plain/black family
20 Islamic Figures	شيخ ازهر شريف مؤسسة علماء قومية دار عين طرق قرضاوي sheikh Azhar Sharif institution ulema nationalism/solidarity house (of Islam) eye path Qaradawi
21 Agriculture & Food Prices	اسعار زراعة قمح سكر لحوم غذائية جنيها جني كيلو انتاج

Topic Label	Ten Most Likely Words
	prices agriculture wheat sugar meat/livestock food pounds reap/harvest kilo production
22 Transportation	نقل سيارات كهرباء محطة طرق محطات طاقة سيارة مهندس مقطورات transfer/transport cars electricity station roads/path stations energy cars (tractor trailers) engineer
23 Corruption	برغم وسوف ثروة مسئولين مسئول رشوة وبرغم سريعة لا يمكن سليمة despite (will be) officials revolution bribe/corruption despite links/rapid (not possible) sound/proper
24 State Land Auctions	اراضي مشروع سكنية جنبي مساحة وحدات اسكان زراعية هيئة فدان land/territory project/plan residential reap/earn area units housing agricultural body/committee acre
25 Internet & News	انترنت قناة شبكة اتصالات قنوات تليفزيون اخبار فيسبوك بوك الالكتروني internet channel network communication channels television news book Facebook electronic
26 Lebanon & Hariri Trial	لبنان حزب حريري محكمة لبنانية لبناني سوريا اغتيال سعد سوري Lebanon party Hariri trial Lebanese Lebanese Syria killing Saad Syrian
27 Election Results	حزب انتخابات دائرة مرشح فئات مرشحين انتخابية مقعد عمال لجان party elections circles/rounds candidate categories/groups candidates electoral seat labor committees
28 Complaints about Government Corruption	هيئة جهاز ادارية تقرير رقابة محافظات مخالفات سيتم اشراف تنسيق body/committee system/apparatus administrative decision censorship/supervision irregularities governorates supervision (to be) complaints
29 Nuclear Iran	ايران ايراني نووية نووي ايرانية اسلحة طهران عقوبات طاقة محادثات Iran Iranian nuclear nuclear Iranian weapons Tehran sanctions energy talks
30 Yemen	يمن صالح بريطانيا بريطانيا بريطاني يمني قاعدة مغرب صنعاء يمنية Yemen Saleh Britain British British Yemeni al-Qaeda (North Africa)/Morocco Sanaa Yemeni
31 Money & Investigations	اموال مستشار نيابة تحقيقات عادلي استيلاء بلاغات سابقين بلاغ طارق money representation investor investigation Adli

Topic Label	Ten Most Likely Words
	takeover/seizure communications previous/former communication/notice Tariq
32 Budget, Spending, & Macroeconomics	دولار مليار جني مليارات اجمالي موازنة اتفاق قطاع بنسبة بلغت dollar million ??? billion total budget spending sector (with respect to) inability/shortage
33 <i>Qur'anic</i> Language	عدل قران جنة صديق سلطان شرف لولا نور وفاء نصر justice <i>Qur'an</i> paradise friend sultan honor (if not for) light/enlighten loyalty/trust victory
34 Constitution & Amendments	مادة دستورية تعديل انتخابات تعديلات شوري مواد مستشار قضاة احكام material/substance constitutional amendment amendments elections consultation material/article judges advisers judgments/provisions
35 Inflation & Rising Prices	اسعار ارتفاع زيادة بنسبة سعر نمو دولار معدل بطالة سلع prices high/rising increase/raise (with respect to) price growth rate/average dollar unemployment goods
36 Egyptian Migrants & Fishermen	سفير سفارة هجرة مصدر مهاجرين جالية مساعد شكر قنصلية مقيمين ambassador embassy migration source immigrants community assistance/aid consulate thankfulness residents
37 Revolution & Protests	متظاهرين ميدان متظاهرون مظاهرة مبارك مسيرة محتجين هتافات شرطة مسيرات demonstrators square/plaza demonstrators demonstration Mubarak march protesters cheers police marches
38 Certificates & Papers	رقم اوراق بيانات بطاقة شهادة جريدة بطاقات تسلّم قيد تسجيل number paper/money/securities data/statements card/ticket certificate/degree cards/tickets newspaper under/restriction receive/receipt/delivered registration
39 Crime	لواء جيزة شرطة مباحث مساعد اكتوبر عامل عميد اسماعيل قبض brigade/banner brief/soon police investigation aide/assistant October labor dean/provost Ismaili catch/seize
40 Stock Market	سوق بورصة نقطة شركات مؤشر مستثمرين تعاملات شراء اجانب بيع market (stock market) point index companies investors transactions buy/purchase foreign sale
41 Frozen Assets	وثائق وثيقة برقية سرية اصول سفارة سفير سويسرا سويسرية تجميد

Topic Label	Ten Most Likely Words
	documents document cable/wire secrecy assets embassy ambassador Switzerland Swiss freeze
42 Pensions, Tax Exemptions, & Payments	مشروع تأمين ضرائب ضريبة غالي اسر خدمة خدمات شركات معاشات project insurance/security taxes taxation expensive families service companies services pensions
43 Sudan	اتفاقية جنوب اببي قبائل قبيلة شمال تصويت ولاية اثيوبيا مسيرية agreement/convention south Abyei tribes tribe north development Ethiopia state Misseriya
44 Copts & Sectarianism	كنيسة اقباط اسكندرية حادث مسيح بابا كنائس قدسين ميلاد تفجير church Copts Alexandria accident/incident Christian Pope churches (Two Saints) Christmas explosion
45 Flooding	امطار ميا جوية درجات فيضانات حرارة سيول رياح تسببت طرق rain water air degrees flooding temperature Seoul wind caused paths/roads
46 Sudanese Independence	جنوب سودان استفتاء شمال انفصال سوداني خرطوم بشير مؤتمر سودانية south Sudan referendum north separation Sudanese Khartoum Bashir conference/meeting Sudanese
47 Jordan & Kuwait	اردن ملك عمان نواب اردنية اردني كويت كويتي نجار ناصر Jordan Amman king parliament Jordanian Jordanian Kuwait Kuwaiti ??? Nasser
48 Poetry-influenced Language	شمس نساء ماء مني امي عين عيني سعادة شاب خير sun women water/wet mona (my mother) eye (his excellence) eye young woman
49 Islam and the <i>Qur'an</i>	كريم رسول وسلم تعالي صلاة صلي خير نبي مسجد اسلام kind/generous messenger peace almighty prayer pray good Prophet mosque Islam
50 Illness & Death (Swine Flu)	وتم صحة مصدر كمال حالات مركزية مصطفى انفلونزا تاكد صلاح (and then) health source integrity cases central Mustafa flu ensure minister
51 European Union & Turkey	اتحاد اوروبي تركيا يورو تركي تركية يونان مساعدات خطة انقاذ union European Turkey Euro Turkey Turkish Greece aid/assistance plan/outline save

Topic Label	Ten Most Likely Words
52 Citizenship & Nationals	قومية عامل اهرام عاجلة دائمة يومي اكتوبر خطة بند سحب nationalism labor/union Ahram urgen permanent daily October plan/outline item/case immediate/prompt
53 Theft, Crime, & Drugs	نيابة تحقيقات واقعة مباحث مستشار محامي تحريات متهمين وتبين مجني representation investigation incident investigation adviser/counselor lawyer inquiry (accused of) shows/finds victim
54 Celebrations	انت احنا عيد اغاني فرحة فرح عشان اغنية دموع لازم you (Ehna TV) holiday/celebration music joy Ashan joy/gladness song/music tears necessary/required
55 Museums, Theft, & Looting	متحف قطع ذهب اثرية سرقة ميني حضارة قطعة حجر حواس museum cut/disconnect gold antique theft building civilization/culture piece/segment stone Hawass
56 African Union	اتحاد افريقيا افريقية افريقي قارة مؤتمر وشمال مغرب اقليم افارقة Africa union African African continent meeting/conference north (North Africa)/Morocco province Africans
57 Governorates	محافظة مركز محافظ قرية اهالي منيا بمحافظة بني اسكندرية كفر governorates center governor village people Minya province Bani Kafr el-Sheikh Alexandria
58 Uninterpretable	جنسية ملكية حالات جدي ممكنة نقطة غموض ارتباط سير اعتراض ownership nationality cases/situations serious possible/potential point ambiguity correlation/connection objection/protest course/progress
59 Election Monitoring	مراقبة انتخابات تغطية سير تعليمات رقابة متابعة لمراقبة انسحاب اغلاق control/monitor elections coverage/concealment course/progress instructions censorship/supervision follow-up (to monitor) withdrawal close/(shut down)
60 Arab League	زيارة لقاء موسي سليمان عمرو تقي لبحث اجتماعا لقاءات زيارت visit meeting Moussa Suleiman Amr pious/devout (to discuss) meeting/together meetings meetings
61 Accidents & Collapses	منازل قرية اطفال انهيار مجاورة منزل مصنع عائلة مكون خيمة

Topic Label	Ten Most Likely Words
	homes village children collapse/fall factory neighborhood home family component tent/pavilion
62 ``Civil Society"	جمعية مركز مؤسسة مراكز جمعيات صحية خدمات احتياجات كفاءة صحة association center institution centers associations health needs services competence health
63 Reader Letters	طعام بلطجية سوداء كرسي نوم بيضاء تلفيز خوفا راحة ملابس food thugs chair/seat plain/black sleep white/blank TV comfort/rest fear clothes
64 Trade & Industry	تجارة صناعة تجارية سلع سوق مصانع رشيد انتاج منتجات صناعات trade industry trade-related goods/merchandise market factories Rashid production products industries
65 Algeria Protests	شرطة جزائر شغب متظاهرين مبني نيران حريق حجارة اشتباكات عيان police Algeria anger demonstrators building fire/flames fire/combustion stone clashes eyewitness
66 Asian	صين جنوبية كوريا شمالية كورية صينية صيني يابان كوري شمالي China southern Korea northern Korean Chinese Chinese Japan Korean north
67 Labor strikes	عاملين عمال اجور شركة عمالة رواتب موظفين قطاع موظف جني workers union/labor wages company labor/employees salaries officials/employers official/employer sector incentives
68 Human rights	تقرير حملة لحقوق انتهاكات اعتداء مركز تعذيب نشطاء خالد عدل investigation campaign rights violations assault center torture established Khaled justice
69 Gaza, Israel, & Palestine	اسرائيل غزة اسرائيلية اسرائيلي قطاع حماس نتناهو صحيفة فلسطينية فلسطين Israel Gaza Israeli Israeli sector Hamas Netanyahu newspaper Palestinian Palestine
70 The Palestine Papers	فلسطينية اسرائيل فلسطين فلسطيني قدس مفاوضات اسرائيلي اسرائيلية عباس استيطان Palestinian Israel Palestine Palestinian Jerusalem talks Israeli Israeli Abbas settlements
71 Book Expos	كتاب معرض شاعر فنية شعر ادب مؤلف فعاليات فنون ترجمة book exhibition poet artistic poetry literature author events arts translation

Topic Label	Ten Most Likely Words
72 America	امريكية اوباما امريكي اميركية كليتون مبارك امريك امريكا وزيرة بوش America Obama American America Clinton Mubarak American America minister/secretary Bush
73 Khaled Saeed Movement	سعيد خالد تيار قصر جوع عار صفحة اسود رسائل جسد Saeed Khaled movement bridge hunger/starvation shame/disgrace page/spread Aswad messages body
74 Elections, Brotherhood, & Baradei	اخوان جماعة حزب مبارك انتخابات برادعي سليمان جبهة جمعية شريف Brotherhood group/association party Mubarak elections Baradei Suleiman front association Sharif
75 Iraq	عراق عراقية عراقي بغداد مالكي غزو مصدر كتل تحالف محافظة Iraq Iraqi Iraqi Bahdah Maliki invasion masses source alliance province
76 American politics	نواب جمهور شيوخ تجديد ديمقراط زميني جمهوري كونغرس استطلاع اقر Congress republic sheikhs renewal Democrat (time frame) republic Congress reconaissance table/schedule
77 Universities	جامعة جامعات علمي علوم كلية علمية طلاب بجامعة دراسة عالي university universities scientific sciences scientific/secular college students university studies high
78 Israel, Palestinian Prisoners	احتلال ابو اسري مشروع مصالحة هدم اقصي فصائل خليل جدار occupation Abu prisoners plans conciliation demolition Aqsa factions Khalil wall
79 Wikileaks	صحيفة اميركية ويكيليكس وثائق سرية بريطانية اسانج تايمز امريك بوقبات newspaper American Wikileaks papers Britain secrecy Assange Times American cables
80 Russia, EU, & Italy	اتحاد اوروبي روسيا روسي روسية ايطالية معاهدة ايطاليا موسكو ايطالي union European Russia Russian Russian Italy treat/pact Italy Moscow Italy
81 Deaths, Woundings, & Accidents	حادث سيارة مستشفى نقل مصابين جثة مصرع ضحايا سيارات سائق incident/accident car hospital transfer/transport patients corpse death cars victims driver
82 Governorates, Inspections, & Demonstrations	محافظة محافظ اسكندرية مرور سيارات لواء عيد محافظات نظافة بمنطقة governorate governor Alexandria passage/traffic cars

Topic Label	Ten Most Likely Words
	brigade/banner celebration governorates cleanliness/purity morning
83 Ivory Coast	ساحل غباغبو عاج وتارا رئيسا ولايت انصار منتهية تنحي حفظ Gbago Sahel Ivory Ouattara president states supporters finished (step down) save/preserve
84 Mubarak and Government Statements	وطن خير شرفاء تخريب سادة تطرف ماحداث حضاري واستقرار خطابات nation/NDP good honest sabotage/vandalism masters (what happened) civilized radicalism stability speeches
85 Football/Soccer	فريق منتخب مباراة بطولة كرة فوز نادي كاس اتحاد لكرة team/squad elected/selected match championship football vicotry club cup union football
86 Spies & Suicide Attempts	ضابط شاب شقيق قاء وطلب ضباط قرر يدعي عمرو انتحار officers youth brother demanded meeting officer claims decided suicide Amr
87 Names	مصطفى ومحمد واحمد سعيد يوسف اهرام فتحي نصر منصور فؤاد Mustafa Muhammad Ahmed Saeed Ahram Yousef Fathi Nasrallah Fouad Mansour
88 Prison	شرطة سجن افراج اعتقال سجون معتقلين ضباط قبض سجناء محاكمة police prison release prisoners arrest detainees officers catch/seize prisoners trials
89 Nile Basin & Water	ميا نيل حوض بيثة صرف متر صحي مائة نهر شرب water Nile basin environment exchange meter healthy water-related river drink
90 Germany and France	المانية مانيا فرنسا فرنسي الماني ساركوزي باريس بحرين وزيرة برلين Germany German France French German Sarkozy Paris Bahrain minister Berlin
91 Mubarak Talks & Meetings	مؤتمر قمة مبارك وفد بلدين مشترك ابو جلسة مباحثات اجتماعات meeting/conference summit Mubarak Wafd/delegation (two countries) joint/common session/meeting Abu talks meetings
92 Courts & Trials	محكمة احكام جلسة دعوي قضائية اداري متهمين برئاسة محاكمة قاضي court courts session/hearing case/claim judicial administrative accused chaired/heard trial judge

Topic Label	Ten Most Likely Words
93 Tunisia	تونس تونسي تونسية ثورة سيدي غنوشي مخلوع بوزيد حزب بطالة Tunis Tunisian Tunisian revolution Sidi Gannouchi deposed Bouzid unemployment party
94 Air Travel (Hajj)	حجاج طيران مطار طائرة رحلات طائرات رحلة سفر بعثة جوية pilgrims plane airport plane trips aircraft trip travel mission aerial
95 Smoking, Obesity, & Children	دراسة اطفال مركز نفسية جسم تناول تدخين مواد اعمار دقيقة study children psychological center body take/address smoking material/substance ages
96 Lawyers Syndicate & Unions	نقابة محامين جمعية نقابات نقيب عمومية صومال مهنية ميزانية مهنة syndicate/guild lawyers association syndicates captain/chair public Somalia professional budget profession
97 South Asia	هند انقلاب استراتيجي انهيار هندي هندية اختراق اندونيسيا شرقية سرعة India coup strategic Indian collapse Indian penetration/breakthrough Indonesia Eastern speed/pace
98 Films & Festivals	مرأة فيلم مهرجان جائزة افلام رواية نساء سينما حفل اخراج woman film festival award films novel women cinema ceremony produced
99 Sinai & Sharm el-Sheikh	بحر سياحة احمر سياحية سويس قناة بحرية سيناء شيخ ميناء sea tourism red tourist Suez canal/channel nautical Sinai port Sheikh
100 Government Reform Statements	مسئولية فئات وطن برامج منظومة قاعدة مجتمعية معالجة مسيرة مجالات responsibility categories/groups nation/NDP parliament system community/societal base/al-Qaeda treatment/handling march areas/fields