A Model of Crowd-Enabled Organization: Theory and Methods for Understanding the Role of Twitter in the Occupy Protests

SHEETAL D. AGARWAL¹
W. LANCE BENNETT
COURTNEY N. JOHNSON
SHAWN WALKER
University of Washington, USA

This analysis establishes a conceptual framework, empirical criteria, and measures for deciding when technology-equipped crowd networks such as Occupy Wall Street behave as organizations. The framework is based on three principles that underlie most organizations: (1) resource mobilization; (2) responsiveness to short-term external conditions; and (3) coordinated long-term change, adaptation, or decline. We argue that Twitter played a coordinating role in Occupy as a connector and dynamic switching mechanism linking various networks. We develop methods for tracking how users embedded and shared links to resource locations. Using a database of some 60 million tweets, we examine different types of links distributed through different hashtags across time, showing how Occupy operated along each theoretical dimension as a networked organization.

Keywords: Occupy, Twitter, networks, networked organization, collective action, big data

Introduction

During a peaceful Occupy protest in Oakland, California, on October 25, 2011, a police-thrown projectile critically injured Scott Olsen, an Iraq veteran and protester. Within one day, a Facebook page

¹ The authors wish to acknowledge the support of National Science Foundation grant 1243170. INSPIRE: Tools, Models and Innovation Platforms for Research on Social Media. We also thank our colleagues Robert Mason, Joe Eckert, and Jeff Hemsley at the Social Media Lab (SoMe Lab http://somelab.net), University of Washington.

Sheetal D. Agarwal: sda25@u.washington.edu
W. Lance Bennett: lbennett@uw.edu
Courtney N. Johnson: cnj3@uw.edu
Shawn Walker: stw3@uw.edu
Date submitted: 2013–02–13

Copyright © 2014 (Sheetal D. Agarwal, W. Lance Bennett, Courtney N. Johnson, & Shawn Walker). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at http://ijoc.org.
was created to draw attention to Olsen's situation and address police brutality. On the same day, the main Occupy Wall Street website, occupywallst.org, posted an update on what had occurred, a YouTube video of the event, and a call for action. The #occupyoakland hashtag spiked in popularity as news spread on Twitter via shared links from news websites, Occupy websites, blog posts, Facebook pages, and YouTube. In the days following, Occupy camps across the country used websites, social networking sites, and on-the-ground meetings to coordinate marches in solidarity with Oakland.

Many such events pulsed through the complex and connected crowds that defined the movement protesting the most serious financial crisis since the Great Depression of the 1930s. Encampments and public demonstrations spread rapidly across the United States and beyond, reaching 952 cities in 82 countries in a matter of months in 2011 and 2012. These complex networked collectivities represent an increasingly common form of organization that challenges previous theory and methods. Our analysis seeks to establish empirical criteria by which such networks of networks perform various kinds of work and can reasonably be called "organizations." We develop a framework to understand such multilayered networks based on three principles (of several) that define most organizations, whether bureaucratic or "postbureaucratic": (1) resource mobilization; (2) responsiveness to short-term external conditions; and (3) coordinated long-term change, adaptation, or decline. The rationale for highlighting these three particular organizational properties is explained below. The advantage of identifying and measuring core organizational properties common to most bureaucratic organizations and some complex networks is to better document and compare different forms of collective action, and in particular to understand the workings of emergent large-scale mobilizations such as Occupy, the indignados in Spain, and the Tahrir Square protests in Egypt. Such actions have become increasingly common, and they are characterized by few conventional bureaucratic organizations at their core (Castells, 2012).

**Occupy as a Postbureaucratic Social Movement**

Occupy encampments and solidarity networks typically emerged from dispersed grassroots initiatives following the model of the New York City protests in September 2011. Membership was generally open to those individuals frustrated by economic and social conditions who were willing to engage in inclusive and egalitarian modes of participation. Occupy adopted the ethos of inclusiveness and diversity established in the global social justice movement of the previous decade. Many issues and styles of personal political expression were welcomed, as long as they did not conflict with the core principles of

---

2 See https://www.facebook.com/weareallscottolsen.
3 "We are all Scott Olsen: Occupy Oakland #OWS" (October 26, 2011), http://occupywallst.org/article/scott_olsen
6 For discussions of postbureaucratic organization, see Bimber (2003), Heckscher (1994), and Hodgson (2004).
7 Despite the ethos of inclusiveness, several surveys have shown the core participants to be generally young, White, above average in income, and left-leaning (Milkman, Luce, & Lewis, 2012).
leaderless democratic consensus adopted as the operating code across the local General Assemblies, which constituted the physical political hubs of the movement. Beyond the local assemblies, camps, and protest events, many more dense communication networks utilized technologies ranging from SMS and e-mail to blogs, websites, Facebook, YouTube, Livestream, and Twitter. We propose to track how such networked organization grew, intersected, and morphed in response to events that ranged from marches to violent police actions and camp evictions. Throughout these developments, one media platform among others—Twitter—emerged as key in flexibly stitching together a dynamic and responsive network of networks. The capacity of Twitter to stitch together different human- and technology-based networks has helped other dispersed populations (such as global disaster relief networks) perform various kinds of “crowd work” (Starbird, Muzny, & Palen, 2012), and we argue the same phenomenon occurred within Occupy.

Crowd-enabled networks are one result of a distinctive “logic of connective action” in which large scale individual engagement is both highly personalized and technology enabled, creating diverse paths for individuals to participate in and activate their own social networks (see Bennett & Segerberg, 2011, 2012, 2013). The logic of connective action is different from conventional assumptions found in theories of collective action that stress the costs of individual participation and the importance of sustaining engagement through common group identification (Bennett & Segerberg, 2013; Olson, 1965). A common organizational problem in conventional movements is how to get people to take risks, share costs, adopt common identities, and join actions rather than wait on the sidelines (McCarthy & Zald, 1973). These movement issues evoke the logic of collective action, which centers on overcoming the “free-rider problem,” which is defined as nonparticipants riding for free on the efforts of participants who bear the costs of action (Olson, 1965). The typical solution to this problem involves hierarchical organization and leadership to mobilize resources such as knowledge, education, training, and various participant benefits aimed at inducing people to join and act together. In this scenario, mobilizing becomes difficult without the support of conventional organizations that can help in coordinating individuals and resources toward a common goal (McCarthy & Zald, 1973, 1977). In addition to providing resources, these organizations develop collective action frames for participants to share common group-based identities. These mechanisms for group solidarity are typically found in familiar issue- and identity-based social movements such as civil rights, peace, feminism, anti-nuclear power, or nongovernmental organization–based environmentalism. We argue that, as some forms of mobilization transition from collective to connective action, they assume hybrid organizational properties that are not easily categorized or explained using conventional models of organizations (Bimber, Flanagan, & Stohl, 2012; Chadwick, 2013). Following these trends, more recent social movement theories have also departed from the early roots in collective action theory, but much of the field remains focused on organizations, coalition formation and brokerage, and collective action framing (Bennett & Segerberg, 2013).

Occupy and other digitally networked organizations have typically been labeled social movements, but they do not fit easily into traditional organization-centered, coalition-based social movement frameworks. The rise of these connective action networks begs for a theoretical bridge between connective and collective action, both to focus empirical work and distinguish among the different organizational principles that may be in play. Many classical issue and identity movements are heavily managed in terms of who is in or out, how the network relations among formal organizations are
structured, and how issues are framed (Snow & Benford, 1992). By contrast, Occupy invited rank-and-file action initiatives, which others could join or not, while maintaining common bonds with an inclusive and easily personalized action frame of “We are the 99%” (see Bennett & Segerberg, 2011, 2012, 2013, for further discussion). As a result, crowd-enabled forms of connective action shared resources and coordinated actions via dispersed face-to-face and digitally assisted communication infrastructure. These technology-enabled networks appeared to be “leaderless” (Syrek, 2012), yet often displayed remarkable levels of coordination while blurring many distinctions between virtual and face-to-face networking (Payne, 2012). The openness to diverse positions and multiple issue frames in crowd-enabled networks raises questions about how crowds actually orchestrate common actions, manage participation, and coordinate goals (Bennett & Segerberg, 2012, 2013; Castells, 2012). In particular, if formal, hierarchical organization is hard to locate in these dense networks of connective action, then what sort of organization are we talking about?

We argue that Twitter played a coordinating and organization-building role in Occupy by acting as an overarching connective and switching mechanism (Castells, 2007), linking different networks. This organizational process was driven by diverse users embedding and sharing different types of links from a broad array of websites and platforms in their tweets. Those links contained an impressive array of resources that flowed in various ways and at particular times through Occupy networks. Following the elements of organizational process outlined above, our empirical studies address three main questions. First, what kinds of linked resources were present in Occupy Twitter streams, and did their production and allocation reflect meaningful patterns? Second, did patterns in the sharing or allocation of those resources reflect systemic responses to short-term events? And, finally, did any patterns of linking, resource sharing, and network traffic direction using Twitter hashtags suggest longer-term evolutionary change in the Occupy network? By examining link types, destinations, and the dynamics of the linked resources through different Twitter hashtags in the U.S. mobilization, we gain insight into how the Occupy crowd functioned as a “crowd-enabled” network—or, more accurately, as a network of networks (Bennett & Segerberg, 2013). “Network of networks” is a concept applied to examinations of complex organizational structure. When an individual actor belongs to more than one network, the individual’s joint membership effectively connects the two networks, creating a larger network (Craven & Wellman, 1973). Understanding Occupy as a network of networks, or multiplex networks, provides insight into how networked individuals and informal groups—such as on-the-ground camps, General Assembly networks, and various digital media networks—connected to create a larger networked organization.

Crowd-Enabled Organization: A Theoretical Framework

Understanding how complex networks emerge and achieve common business, political, or social goals is an area of interest across various theories and fields, including organizational communication, communication network theory, international relations, political economy, governance, and collective action theory. Our model of networked organization builds on prior work in these fields. Contributors to

---

8 Our analysis does not privilege Twitter for all types of crowd organization. Other cases may be integrated through other social media, and Twitter may be replaced by other technologies in the future. However, our general model should translate to these other cases.
this literature focus on properties such as boundaries, intentionality of purpose, or recognized membership as necessary components of networked organizations. We challenge and supplement these assumptions by suggesting that much less clearly defined networks such as Occupy can also generate networked organization without these structural properties.

In organizational studies of global businesses, “networked organization” is a term used to describe the nature and attributes of what are often termed “boundaryless organizations” or virtual organizations (Ahuja & Carley, 1999; Monge & Fulk, 1999). Structure in these networks is provided by formal organizations that establish network linkages such as strategic alliances, collaborative pacts, and joint ventures for the purposes of achieving collective goals (Powell, 1990). These loose business networks can operate without market/hierarchy models based on formal agreements and exchange mechanisms; instead, they often cooperate or compete on the basis of far looser relationships that involve trust.

Moving beyond business, Ostrom’s (1990) theorizing on collective action identified the emergence of collectivities to manage common pool resources such as water and fisheries. She noted the frequent requirements of bounded community membership, capacities for monitoring and policing member behavior, shared definitions of the common issues, and intentional organization and design of an organized governance network. By contrast, examples of crowd-enabled organization often lack a formal bureaucracy, formally recognized hierarchical authority relationships, and other traditional structuring conditions (Mueller, 2010; Rheingold, 2002). The goal of our investigation is to assess whether associational patterns established through communication technologies create similar structures and processes found in more conventional organizations.

Mueller (2010) suggests an interesting theoretical move in his work on networked transnational policy and governance organizations. He observes that Powell’s idea of trusted relationships as an organizational principle resonates with assumptions about peer production in the field of digital media (e.g., Benkler, 2006; Castells, 2012; Raymond, 2001; Rheingold, 2002; Shirky, 2008). At the heart of this line of thought is Benkler’s idea that technology enables loose ties that replace trust based on direct personal relationships.

Yet the sort of network Mueller is interested in understanding (the global Internet governance regime) more closely resembles politically brokered, issue-focused networks (Keck & Sikkink, 1998). In Mueller’s scheme, network organizations are intentional, self-organizing groups, generally composed of identifiable constituent organizations, and consciously organized on the norms of reciprocity. Importantly, actors in his definition of network organization display agency and intentionality, in part by establishing criteria for deciding membership. Thus, despite making tangential contact with peer-produced technology-enabled networks, Mueller’s definition of networked organizations also moves the construct of networked organization in bounded and intentional directions.

However, Mueller (2010) acknowledges the existence of many collectivities that are relatively unbounded, unintentional, and not issue- or policy-focused. Since his interest lies elsewhere, he does not theorize these unbounded collectivities that he categorizes as “associative clusters” created by actors engaged in repeated interaction with few clear rules to identify who is included or excluded. The
implication is that associative clusters may lack the focus or decision-making capacity of more formal networked organizations such as his transnational governance regimes.

We are not convinced that the organizational standing of crowd-enabled networks should be settled by the definitional choices of scholars examining other kinds of cases. Rather, we seek to understand what sorts of organization may exist in collectivities with relatively unspecified geographic or membership boundaries, that lack singular purpose or defining issue focus, and that display little formal hierarchical organization. Rather than reifying particular definitional claims, we ask whether complex networks can perform various tasks and display processes commonly associated with other kinds of organizations.

In the broad literatures on forms of bureaucratic and postbureaucratic networked organization, we have identified three structuring properties that seem important for determining whether there is coherent organization in technology-enabled crowds and for comparing the performance of those crowds to other kinds of organization. For example, in regard to our first element of resource mobilization and distribution, Ostrom (1990) emphasizes resource allocation in her focus on how common pool resources, which include natural or manmade resources such as forests, oil fields, and fisheries, are managed. Organizational studies give attention to identifying types of resources, such as information (Powell, 1990), skills, knowledge, labor, and finances (Ahuja & Carley, 1999), and how they are exchanged and allocated across and within firms. Resource management is a key category Ahuja and Carley (1999) identify as task types critical to the networked organization. Mueller (2010) examines on how state and nonstate actors manage critical Internet resources such as IP addresses and domain names.

Our second defining property, responsiveness to external events, is also a phenomenon addressed by many students of both bureaucratic and networked organization. Among the classic works in organization theory are the theories of organizational learning by James G. March and colleagues (e.g., Cyert & March, 1963; March & Olsen, 1973). These approaches address both short-term responsiveness, our second organizational criterion, and longer-term adaptation patterns, our third defining element. March (1981) notes that short-term responsiveness typically results from factors such as sudden problems that engage decision processes; conflicts triggered by outside forces or breakdowns within the organization; and contagions that spread new action repertoires, problems, or conflicts (March, 1981). In studies of less bounded organizational forms, Ahuja and Carley (1999), for example, point to the identification of problems (e.g., bugs in software) and responses to those problems (e.g., software patches) as important to collective maintenance of virtual organizations. Ostrom (1990) also illustrated the network responses to resource problems such as drought and overfishing, while Powell (1990) noted the ability for networked organizations to respond more quickly and efficiently than hierarchical organizations to unanticipated changes and market fluctuations. In his evaluation of “smart mobs,” Rheingold (2002) identifies the ability of peer-to-peer networks to use technology to organize and respond to changes in the environment as a critical component of the underlying logic of this type of organizing.

Creating grounds for making such comparisons is important, particularly in light of the frequent dismissals of crowds as not coherently organized, as weak politically, or as products of idle “clicktivism” (see, e.g., Gladwell, 2010; Morozov, 2011).
Our third defining element is long-term coordinated adaptation and response to environmental changes. As with the other two elements of organization, adaptation and decline occupy a rich literature in classical studies of bureaucratic organization (see, e.g., Levinthal, 1991; March, 1981). Adaptive processes are also important in Ostrom's (1990) work as she identifies how the creation of different governing documents and changes in community responses shape the resource allocation process and procedures over time. Organizational studies evaluate the flexibility of network organizations that allow interorganizational linkages to wax and wane over time to accommodate unforeseen changes over time (see Monge & Fulk, 1999). And preliminary evidence suggests that crowds change their communication patterns in systematic ways as they go into decline and seek new organizational forms or more stable relationships (Bennett & Segerberg, 2011).

These elements appear explicitly or implicitly in all of the above theories, and they appear whether the organization form is bureaucratic and hierarchical or loosely networked and postbureaucratic. They appear whether the theoretical focus is on membership rules, boundary maintenance, decision making, or other features that may distinguish one organization type from another. These elements can also be traced back to root models of organizations in environments of imperfect certainty, in which rationality and control are "bounded" or limited by conditions of information scarcity, environmental instability, and other factors that create more open systems than were assumed by earlier command-and-control theories of organizations (Thompson, 2003). We therefore argue that organizations of varying types perform these few tasks in repeated or routinized fashion:

1. Resource mobilization and allocation: producing, gathering, and allocating the material and symbolic goods that (are thought to) better enable the collectivity to operate.
2. Responsiveness to short-term external conditions: recognizing near-term threats and opportunities and adjusting responses in concerted fashion.
3. Coordinated long-term adaptation, change or decline: reflected in systematic internal shifts in the production and deployment of resources.

Our project seeks to establish criteria for deciding whether crowds such as Occupy display these coherent organizational properties. The initial test of these organizational behaviors relies on analysis of a large trove of Occupy Twitter data that enabled us to categorize, code, and track the resources and flow patterns contained in links. Our study addresses the following research questions, with each question linked to the identification of the corresponding organizational element:

RQ1: (resource allocation): Do different types of link resources flow in different proportions in different Twitter hashtag streams? If so, do those patterns suggest different organizational roles for streams operating at different levels of the crowd, such as those carrying more local or national traffic?
Internatioanl Journal of Communication 8 (2014) Role of Twitter in the Occupy Protests 653

RQ2: (responsiveness to external events): Are there changes in resource link patterns at different points in time and for different hashtags? If so, do those changes reflect organized responses to identifiable short-term events?

RQ3: (longer-term adaptation and decline): Are there notable changes in link and hashtag patterns in the tail of the Occupy tweet distribution? If so, do those changes suggest longer-term reconfiguration of information signals aimed at redirecting attention and relationships in the network itself?

Because information is a type of resource (Powell, 1990), we argue that links embedded in tweets constitute resources. The tweeted links contained information that Occupiers considered important or relevant to the movement. Links to news articles or political commentary websites provided information regarding, for example, how the mainstream media was currently characterizing Occupy’s protest efforts—useful information for protesters looking to gain support for their movement. Links to personal blogs often included an individual’s perspective on an Occupy-related news article or recent protest event and provided firsthand opinions and accounts likely absent from mainstream news organizations. Tweeting a link to a politician’s website might also have provided vital information by, for example, alerting other Occupiers to official channels by which they might voice their concerns.

By monitoring the distributions and dynamics in hashtag flows containing various resource links, we assess the degree to which resource flows were organized in terms of: (1) being produced and patterned differently in different sectors of the crowd; (2) being responsive to external events impacting the crowd; and (3) changing over time due to broader conditions of resource depletion, production fatigue, loss of interest or attention, internal fraction, or shared perception that a particular action has run its course.

Technology in Crowd-Enabled Organization

Occupy was interconnected through multiple platforms and technologies. To contextualize this study, we observed numerous Occupy networks in real time to document some of the scale and diversity of network nodes. Between October 2011 and March 2012, we identified more than 450 distinct Occupy Facebook pages. In addition, we found more than 250 Occupy-related local and national websites, 179 Livestream feeds, more than 2,500 Meetup groups, and more than 355 Occupy-specific hashtags. Links to these and other Occupy networks were regularly shared in tweets during our data collection.

Among these many platforms, Twitter in particular emerged as a tool enabling communication between activists on the ground, creating a global community of geographically dispersed sympathizers who distributed information and created an attentive public for the protests (Howard et al., 2011). We think about Twitter as an overarching organizational mechanism in the Occupy networks based on examining many views of the hyperlink ecology of the multitude of Occupy sites. To confirm, we used IssueCrawler, a tool that examines links and network properties among websites (Bennett & Segerberg, 2013; Bruns, 2007; Devereaux, Cukier, Ryan, & Thomlison, 2009; McNally, 2005; Rogers, 2012), which
revealed how networks of Occupy sites linked to each other over time. We initiated network crawls from different starting points such as technology developer websites, city websites, bridging sites (e.g., interoccupy.net, occupytogether.org), and campaign (“Occupy our homes”) or event (general strike) sites. In all of these crawls, Twitter consistently emerged as the most linked-to site, or the largest and most central node in every network.

More than a thousand personal and camp-based Occupy user accounts emerged, and many thousands more ordinary Twitter users joined in shaping the networks along which resources flowed. Twitter allowed participants to act in a collective movement while maintaining their individuality and various qualities of personalized engagement that seem central to understanding connective action (Bennett & Segerberg, 2011, 2012, 2013). Even more interesting for our purposes was that people in the crowd created and shared new hashtags as a means of directing resources across the widely scattered collection of virtual and physical nodes (Starbird et al., 2012).

Figure 1 offers a visualization of how we imagine Occupy and similar crowd-enabled organizations as a network of networks. The left view simulates how three network layers might be visualized in a top view. The center scheme shows identifiable layers that make up the larger network (e.g., these networks might be the interpersonal networks in camps, the national network of city websites, Facebook networks, telephone trees, interlinked event websites, etc.). The right view shows how these layered networks can be linked through dynamic connective mechanisms such as Twitter, which is represented by the threads establishing connections from one network to another. These threaded connections in our model are established by the links to organizations or sites in different networks that are inserted in more than half of all the tweets and retweets in our database.

These connecting threads may act as “switches” in Castells (2007) model of network power, as people work to shift the focus of the crowd. These switching processes are accomplished by using hashtags to direct resources into different layers of the Occupy Twitter stream. Thus, Figure 1 represents a simple abstraction of layered networks in a kind of stop-action that would include other threaded switching patterns as the visualization moves through time and across networks.

---

10 IssueCrawler (www.issuecrawler.net) is a web analysis tool developed at govcom.org by Richard Rogers of the University of Amsterdam.
We next look at how those dynamic Twitter connections operate, and whether they enable organizational functionalities that match our three minimum defining conditions of networked organizational process. This analysis involves tracking and comparing different Twitter hashtags as they send different kinds of resources contained in embedded links through the crowd at different points in time and in response to different kinds of events.

**Methods**

Other observers have documented how Twitter acts as connective strings across social media platforms, with hashtags binding together communities of interest. Hashtags can act as mechanisms for various kinds of network production, including: finding information of interest; weeding out unnecessary information (Huang, Thornton, & Efthimiadis, 2010); sharing information (Java, Song, Finin, & Tseng, 2007), particularly about events (Starbird et al., 2012); injecting messages into existing topical information streams (Huang et al., 2010); and organizing information (Chang, 2010). Users report that hashtags reflect their desire to connect with like-minded individuals (Java et al., 2007) or to be part of a discussion (Letierce, Passant, Decker, & Breslin, 2010). Tagging a tweet can be at
once a community-seeking mechanism (Laniado & Mika, 2010) as well as a tool to enhance already existing networks (Letierce et al., 2010).

Combining Occupy-related hashtags and links within tweets was a common practice (Green, 2011). For example, the hashtag #ows appeared more than 3.6 million times between October 19 and December 31, 2011, making it the most common hashtag during the time frame of our study. This hashtag often appeared in several hundred thousand tweets per day. Links to websites containing various kinds of information and other resources were included in more than half of those #ows tweets. Tracing the patterns of hashtags and links helps us to understand whether (and how) dispersed groups of activists formed a multilayered network organization, and how that organization compares with different kinds of network organizations previously studied by other scholars. Above all, classifying and tracking the resource links in tweets allows us to determine whether communication-based organizational processes enabled the crowd to distribute resources, respond to events, and change systematically over time.

**Empirically Assessing Twitter as a Network Connective Mechanism**

Using our theoretical framework, we explore how participants in Occupy stitched together online and off-line networks through Twitter. Following (Bennett & Segerberg, 2011), we propose using links and hashtags embedded in Occupy-related tweets as both concrete networking mechanisms and a broader window into the resource ecology of the network system. We explore the dynamics of Twitter as an organizing mechanism using a large data set of more than 60 million tweets captured by the University of Washington Social Media Lab (SoMe Lab; somelab.net) from October 2011 through June 2012. For this analysis, we drew primarily from a subset of 20 million tweets gathered between October 19 and December 31, 2011, a peak period of protest activity before, during, and after police actions that broke up the camps.

Our primary analyses focus on tweets containing the hashtags #ows, #occupyseattle, and #occupyoakland during this time period, and continuing to observe the behavior of the tail of #ows through the spring of 2012. We created a research design to explore how those three hashtags—two local, one national—shared resources, how they responded to events that impacted them at different times, and how they behaved in more routine fashion during times without external threats or local events. The rationale for selecting the two local hashtags (#occupyoakland and #occupyseattle) is that both cities had active camps, and, conveniently for our natural experiment purposes, they experienced external events (police evictions and militant protests) at different times. In addition, we conducted ethnographic observations and interviews in Seattle to contextualize the data trends. The most prominent national stream (that often collected and distributed resources from other streams) was #ows, which also served as a hybrid local feed for the New York City protests, registering an intense activity spike surrounding the police eviction of the New York Zuccotti Park camp. Thus, we have a before-during-after natural experiment design for three hashtags that registered similar events (i.e., police raids) in different parts of the network, with added variation provided by the different periods in which these high-impact external events occurred. These different data sets built around hashtags and their link patterns enable us to answer our research questions.
Some of our analyses draw from the whole data set to examine overall activity levels, spikes in different streams, and patterns of links in tweets and retweets. Other analyses require more fine-grained measures of crowd activity, such as what kinds of resources people were exchanging through links and how those link patterns changed across different streams and within the same stream at different points in time. For these analyses, we drew samples of tweets containing links for each of the three hashtags and developed a coding scheme to distinguish among the kinds of organizational resources represented by different types of links.

Before we describe the methods used to collect, process, and analyze the data, we present a conceptual description of what the data and methods represent. We examine the communication patterns of Occupy through the lens of public Twitter activity. To capture this data, we developed a way to identify the Occupy networks and examine the content of the messages flowing within each network. To identify the Occupy network, we use the presence of structural elements—the hashtags #ows (national), #occupyoakland (Oakland), or #occupyseattle (Seattle)—as a rough filtering mechanism. Next we examine the content of messages flowing within these three connected Occupy networks. As described earlier, we identify links as information resources. We see these URLs as extending the content of the 140-character tweet. Finally, we coded the destination of these URLs by opening the links. By applying the URL coding back to the larger data set, we are able to see communication patterns within each of the Occupy networks we examined. Conceptually, the hashtag acts as a network filtering mechanism, and the URLs represent the resources flowing within the network.

**Data Collection Strategy From Twitter**

As noted above, the SoMe Lab archive of Occupy-related tweets from October 19, 2011, to December 31, 2011, contained roughly 20 million tweets (20,645,921). Tweets were collected using Twitter’s Streaming API, which returns tweets matching any of the search keywords occurring in the text, hashtags, @mentions, or URLs within a tweet. A panel of faculty and graduate students curated a list of popular hashtags, keywords, and Occupy city accounts related to the Occupy movement. The resulting data stream was examined at regular intervals for emerging terms. New terms were added to the keyword list after being reviewed by the entire research team, resulting in a dynamic archive based on a list of 355 keywords as data collection continued through the summer of 2012.

Processing of the tweets included expanding shortened URLs, adding metadata to each tweet to make searching and sampling easier. The metadata added to each tweet included:

---

11 Some amount of noise is inevitable when collecting data from social media sites due to the nature of keyword searches. For example, the search term “strike” may also include tweets related to baseball. However, our sampling method used hashtags to signal relevance to Occupy. We found tweets with Occupy-related hashtags maintained a low rate of noise (under 13%). Tweets without Occupy hashtags had a much higher error rate.
● expanded shortened URLs;\textsuperscript{12}
● list of hashtags;
● list of mentions in the tweet;
● number of hashtags, URLs, and mentions; and
● list of data collection keywords that matched within the tweet and the location of the match (hashtag, @-mention, text, or URL) to explain why each tweet was included.

**Sampling From the Large Occupy Tweet Archive**

Because resource allocation and crowd response patterns can be better assessed by looking at what resources flowed through links in the larger data set, we took random samples over time from the archive and coded them for categories of links as described below. Coding files were generated for each of the three hashtags (#ows, #occupyoakland, and #occupyseattle) by randomly sampling 225 non-retweeted tweets with URLs from every other day starting on October 19, 2011.\textsuperscript{13} Random samples were generated for each day via the following process, which ensured each daily sample contained unique links that could subsequently be tracked through the entire data set:

1. All non-retweeted tweets with at least one URL and the appropriate hashtag were extracted from the archive.
2. The tweets extracted from the archive in Step 1 were randomly sorted.
3. The random tweets were sequentially examined to determine whether the URL it contained was part of a prior sample. If the URL was unique across all the samples that had already been generated, it was added to the coding sample and randomly assigned to one of three coders.
4. This process was repeated until the coding sample for the day contained 225 URLs, or the entire sample of tweets with the date and hashtag from the archive had been exhausted. A total of 30,675 links in tweets were coded for the three hashtags reported in this analysis.

Once the random sample of URLs had been created, coding files were automatically generated for each of the coders, who were trained and supervised by the authors. The coding file contained the tweet’s time stamp, the numerical ID of the tweet, the hashtag, the text of the tweet, and the URL to be coded.

\textsuperscript{12} Common Twitter practice is to use URL-shortening services, which create a shortened URL forwarding to a longer URL.

\textsuperscript{13} The SoMe Lab data set begins on October 19. Because the focus is on whether we observe organizational behavior and response to external events over time, this starting point for the data is not an issue for the analyses reported here. The first major events in the Oakland stream occurred after this date in October, the national evictions in New York City were later in November, and Seattle faced camp evictions in December.
Coding the Links: Codebook Development

The codebook (available on request from the authors) comprised 12 mutually exclusive categories of links deemed to contain distinct types of resources that might be of interest for sharing among members of the crowd: (A) basic news, from The New York Times and CNN to The Guardian and Huffington Post; (B) point of view, political commentary, and activist media; (C) specialty sites that go beyond politics, such as Gawker or Kickstarter; (D) conventional political organizations such as parties, government institutions, or officials; (E) activist organizations, nongovernmental organizations, policy organizations, nonprofits, labor and trade unions, and think tanks not directly defined as Occupy, with subcategories for (E1) commercial media sites for these organizations (e.g., Facebook or Instagram) and (E2) custom-built sites that do not use commercially available platforms; (F) Occupy websites or organizations, with subcategories for (F1) commercial media sites for Occupy and (F2) custom-built sites that do no use commercially available platforms; (G) individual, personal websites, including blogs, flickr or YouTube pages; (H) music videos or celebrity websites; (I) other (including foreign-language tweets and spam); and (J) broken links. For the purposes of some analyses in this article, we combine categories E1 and E2 into one category (E) for all activist organization and nongovernmental organization sites. We also combine categories F1 and F2 into one category (F) containing all Occupy sites.

Intercoder Reliability

Three coders conducted this analysis. Results from our intercoder reliability across all categories were high (Krippendorff’s $\alpha = 0.851$, Fleiss’ $\kappa = 0.851$). The average pairwise percent agreement among coders was 86.3%, and the average pairwise Cohen’s $\kappa$ was 0.829. Overall, item-by-item intercoder agreement was also high. (More detailed coding information is available from the authors.)

Results: Patterns that Reflect Three Elements of Crowd Organization

The data on the link types contained in the three hashtags reported in this article come from coding samples of 30,675 tweets containing links: 5,586 for #occupyoakland and 2,682 for #occupyseattle from October 19 to December 31, 2011, and 22,407 for #ows from October 19, 2011, to April 30, 2012. The sample sizes roughly mirror the volume differences of the different hashtag streams in the big data set.

We find evidence of resource allocation, responsiveness to events, and long-term coordinated adaptation in our analysis of link destinations and Twitter hashtag activity. Before looking at the three dimensions of crowd organization, we offer a broad overview in Table 1 of the patterns of links for the three hashtags, including the retweets and links in retweets contained in each stream. We will return to this table in our analyses of crowd organization patterns. For now, it establishes the dominance of #ows as the most heavily used hashtag (with both local New York City and national followings). The three uneven time breaks are useful for noticing dynamics during periods of different kinds of activity in different locations. For example, #occupyoakland spiked during the first time period, when local protesters experienced raids on camps, large demonstrations, and confrontations with city officials. The middle time period captures the police raids on the original Zuccotti Park camp, which was both a watershed moment...
for the protests and for the Twitter crowd. The third time period included some of the most dramatic confrontations in Seattle, including the clearing of the camp and the shutdown of the port.

**Table 1. Total Tweets and Retweets and Links During Different Time Periods, by Hashtag.**

<table>
<thead>
<tr>
<th>Date</th>
<th>Tweets Total</th>
<th>With URLs</th>
<th>Retweets Total</th>
<th>With URLs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#ows</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 19 to November 14</td>
<td>1,504,643</td>
<td>816,276</td>
<td>54%</td>
<td>669,897</td>
</tr>
<tr>
<td>November 15 to November 18</td>
<td>792,623</td>
<td>312,676</td>
<td>39%</td>
<td>428,671</td>
</tr>
<tr>
<td>November 19 to December 31</td>
<td>1,496,575</td>
<td>849,264</td>
<td>57%</td>
<td>643,762</td>
</tr>
<tr>
<td>Total Tweets</td>
<td>3,793,841</td>
<td>1,978,216</td>
<td></td>
<td>1,742,330</td>
</tr>
<tr>
<td>#occupyoakland</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 19 to November 14</td>
<td>345,381</td>
<td>149,343</td>
<td>43%</td>
<td>202,800</td>
</tr>
<tr>
<td>November 15 to November 18</td>
<td>21,025</td>
<td>9,193</td>
<td>44%</td>
<td>11,192</td>
</tr>
<tr>
<td>November 19 to December 31</td>
<td>85,099</td>
<td>39,464</td>
<td>46%</td>
<td>45,536</td>
</tr>
<tr>
<td>Total Tweets</td>
<td>451,505</td>
<td>198,000</td>
<td></td>
<td>259,528</td>
</tr>
<tr>
<td>#occupyseattle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 19 to November 14</td>
<td>19,254</td>
<td>8,719</td>
<td>45%</td>
<td>7,830</td>
</tr>
<tr>
<td>November 15 to November 18</td>
<td>11,773</td>
<td>5,663</td>
<td>48%</td>
<td>6,560</td>
</tr>
<tr>
<td>November 19 to December 31</td>
<td>25,936</td>
<td>11,874</td>
<td>46%</td>
<td>13,479</td>
</tr>
<tr>
<td>Total Tweets</td>
<td>56,963</td>
<td>26,256</td>
<td></td>
<td>27,869</td>
</tr>
</tbody>
</table>

**Resource Mobilization and Allocation as First Principles of Crowd Organization**

As shown in Table 1, activity levels and linking patterns differed considerably, both between and within hashtags at different times. What is immediately clear from the raw numbers shown in Table 1 is that the #ows stream sent out a higher percentage of tweets with links than the two local streams. This suggests that different streams operating at different levels of the crowd perform different kinds of organizational work. The #ows stream was the most used national tag throughout the protest period studied here, and it also became something of a news feed for the New York Zuccotti Park occupation. As
such, it provides an interesting look at how national organization was achieved in terms of creating more information resources than the other more local streams (suggesting evidence for our first key element of organization). During the November 15–18 time period, when the Zuccotti Park camp was raided and evicted by police, it then mirrored the local streams (a point to which we return in discussing the evidence for our second element of organizational responsiveness).

For now, the main point we draw from the data presented in Table 1 is that organizational resources are differentiated both over time and at different levels of the crowd, from more national (with the noted exception of the Zuccotti Park eviction) to more local levels of the crowd. For most of the protests, the prominence, volume, and overarching role of the #ows stream made it the most likely place to send resources out to the most general level of the crowd. Other patterns indicate that the traffic direction and resource insertion activity in the Occupy crowd were far from random. Indeed, as noted, local streams tended to contain fewer links and focused more on daily logistics and rallying supporters. This pattern is consistent with the findings of Conover, Davis, Ferrara, McKelvey, and Menczer (2013), who analyzed the content of tweets based on geolocation and found that tweets that crossed state borders (e.g., by retweeting) were far more likely to contain symbolically rich framing content than tweets that remained in particular locations, which were primarily logistical and rallying in content. Consistent with its central coordinating role, #ows was the most link-rich stream of the three, with roughly 10% more of its tweets containing links compared to the others. It was also consistently the most used hashtag in our every-other-day samples of tweets across the period of the study. In addition, the overall use of Twitter varied in the three streams; the #ows stream remained active across the entire time period, while Oakland activity dropped off dramatically following the eviction of its camp in October, and Seattle activity picked up as its camp eviction neared in December. Table 1 presents clear evidence of linked resources flowing in different volumes and with different content richness at different levels of Occupy networks. To investigate this resource allocation more closely, we next examine the types of content flowing in these different streams.

Figure 2 provides a closer look at types of resources flowing in these streams. Sites run by individuals sharing personalized information (Category G) comprised the largest portion of links in all the streams. As guided by our first pair of related research questions, it is clear that the types of resources circulating in the Occupy crowd balanced highly personal accounts of the protests with more mainstream media accounts, and that this balance shifted heavily toward the personal as the hashtags became more local. Personal sites made up 44% of the Oakland and 36% of the Seattle links, compared with 27% of the #ows links. As we move from the local to the national (#ows), personal websites give way to general news and commentary about activities of Occupy as a whole, suggesting that users in the crowd clearly understood how to use different hashtags for allocating different kinds of content.
Figure 2. Link resource categories in each hashtag, October 19, 2011 to December 31, 2011.
Basic news sites were the second most common resource category within each hashtag stream, making up 21% of #ows, 22% of #occupyseattle, and 16% of #occupyoakland. However, as we note below, the relative differences in the proportions of different content in different streams suggest different kinds of work being done at different levels of crowd integration. It is also noteworthy that tweets rarely linked to government or politicians’ websites (about 1%) or celebrity sites (less than 1%). Non-Occupy activist and nonprofit groups were linked to infrequently as well, receiving 4% to 7% of the links across the three hashtags. This suggests the relatively low importance of formal political organizations, even sympathetic and supportive ones, to members of the crowd. Occupy’s own websites, including commercial social media sites and Occupy-owned domains, comprised 5% to 8% of the links. However, as shown below in our assessment of long-term adaptation and change, the links to Occupy sites increased dramatically as the #ows hashtag trailed off in the early months of 2012, suggesting that remaining activists were pointing to home-base organizations to regroup as the overarching networks began to break down.

**Responsiveness to External Events as a Second Element of Crowd Organization**

It is clear from the data presented in Table 1 that patterns of links in tweets differed in the #ows stream over different periods of time. Perhaps the most critical period in the entire protest was the eviction on November 15, 2011, of the first camp in New York City’s Zuccotti Park, where the Occupy Wall Street protests began. This event was not only highly significant from a symbolic standpoint but also continued a trend in which public authorities across the nation tried to shut down the protests. The number of daily tweets in the #ows stream increased during the New York City camp eviction, while the proportion of tweets containing links decreased substantially. This suggests that this important hashtag was capable of switching from general resource distribution to more direct information sharing and coordination among protesters during a time of extreme conflict with authorities. This is consistent with the findings of Conover et al. (2013) that streams anchored in local traffic tend to be more logistical and less rich in broader protest framing content. The dip in the #ows stream during the Zuccotti Park evictions in mid-November suggests that this stream served as both a national and local tag. The New York City protests were the origins of the movement and the symbolic hub of the communication networks. This means that during most of the peak protest period the #ows stream served as the main content generator, including the highest levels of links and retweeted links (particularly links to national news and political commentary). The exception to this pattern was when the Zuccotti camp was raided in November, at which point the #ows stream responded with proportionately fewer links and retweeted links, more resembling the two local streams across the entire period.
Figure 3. Three Twitter streams showing volumes and spikes in response to event impacts. Dotted vertical lines represent the break points between time periods. Note differences in scale of the y axes. Legend (all dates 2011):

1. October 25: Occupy Wall Street national solidarity with Oakland camp evictions and injury of Iraq War veteran Scott Olsen by police
2. November 15: Police eviction of New York City Zuccotti Park original Occupy camp
3. November 17: International day of action, 30,000 demonstrate in New York City and cities across the nation, many arrests
4. October 25: Police clear Oakland camp, injure Iraq War veteran Scott Olsen
5. November 2: General strike in Oakland, 100,000 march on Port of Oakland
6. November 14: Hundreds of police clear main Oakland camp for the second time
7. November 2: Occupation of Chase Bank, arrests, clash with crowd, use of pepper spray
8. November 15: March through downtown Seattle, police clash, pepper spray 82-year-old woman
9. December 12: March on Port of Seattle, shut down operations, arrests, sparks opposition from unions
As shown in Figure 3, the three streams were activated differently in response to events. We identified key events that occurred nationally and in the New York City, Oakland, and Seattle camps. These events clearly had different effects on different Twitter hashtags. For example, the #ows stream had a small spike following the police eviction of the Oakland camp and an injury to a protester. The same event produced one of the largest spikes in #occupyoakland. The largest spikes in #ows were in response to the police eviction of the original Zuccotti Park camp, and to a national day of protest involving large demonstrations in New York City and across the nation. These activity spikes (ranging from 100,000 to 300,000 tweets daily) suggest that #ows was used as the main national hashtag, with greater attention to the New York City protest hub during the Zuccotti Park eviction period in November, illustrating how this hashtag shifted between local and national organizational roles in response to different events. By contrast, the #occupyoakland and #occupyseattle activity levels were driven by both local and national events, with large spikes surrounding the New York City camp evictions and the related national solidarity protests. What is interesting is that the local spikes in Oakland and Seattle at the time of the New York City evictions did not entail increases in links or retweeted links, providing further confirmation that local activities such as solidarity marches focused more on direct communication among protesters (e.g., logistics and rallying) than in sending more content-rich links and retweeted links. The distinction between local and national events is not always easy to draw, however, as when the New York solidarity-themed Seattle march was soon buried in both the local media and activist communication by police attacks on the demonstrators that included pepper-spraying an 82-year-old woman who had a long history of local public service.

**Coordinated Adaptation Over Time as a Third Element of Crowd Organization**

Figure 3 reveals an even longer-term adaptive response from the crowd: the growth over time (steadily since October, and more rapidly after January 2012) of links to Occupy organizations, suggesting a refocusing of organization around Occupy websites. Recall that we extended our coding of #ows beyond 2011 and into the spring of 2012. As shown in Figure 4, Occupy organizations (e.g., city websites, Facebook pages, Meetups, etc.) were the primary growth resource, even as the number of tweets decreased dramatically in the aftermath of the breakup of camps and the related weakening of face-to-face assemblies. Indeed, more than 20% of all link resources as the Twitter stream diminished pointed to various Occupy sites by the spring of 2012, surpassing all other resource types, with the exception of personal sites.
This adaptive capacity of shifting attention of crowds toward Occupy websites seems a long-term response to various forces such as camp evictions that produced disarray in the protests organizations. Although few conventional brick-and-mortar associations could be relied on for purposes of regrouping, many local and national organizing websites and Facebook groups had been created by late fall 2011, when the camps were displaced. Those virtual organizational sites became the most shared resources in the Twitter streams as the crowds diminished. Similar resource seeking and stabilizing activity has been observed in other protest crowds (Bennett & Segerberg, 2011). Such organizational responses may vary from crowd to crowd, depending on external conditions such as levels of repression, available political opportunities, continuing engagement of bystander publics, and the attitudes of protesters themselves about proper forms of future organization. These contingencies affecting the sustainability and forms of crowd organization are addressed in the conclusion.
Conclusion

The organizational dynamics of crowds offers a rich area for theory and research. Many scholars have begun to examine aspects of coordination involved in certain kinds of activity, such as how particular conversations are organized (Huang et al., 2010) or how bridges across different crowd sectors are constructed to get particular kinds of work done (Starbird et al., 2012). However, more general models of crowd-enabled organization that facilitate empirical comparisons across different crowds and between crowds and other kinds of organization have not kept pace with growing numbers of micro-analyses. We offer our simple model as a start for more general theory building.

Many questions are left unanswered in our analysis. Future analyses could examine the properties and the roles of multi-resource producers. Roughly 10% of the users in our sample tweeted multiple types of links, suggesting that some members of the crowd were more inclined to multitasking when it came to resource identification and sharing. In addition, future analyses may establish how different link categories travel through the larger crowd depending on the hashtags attached to them, the external event contexts affecting their relevance over time, and even the kinds of coordination involved among user subnetworks in promoting particular resources.

Beyond such specific organizational questions, other larger matters of the effectiveness of crowds exist: whether they have an impact and how they can sustain themselves. Of course, these questions also apply to social movements more broadly, and scholars debate how such outcomes compare to more formal organizations (see Bennett & Segerberg, 2012, 2013). Many observers of crowds take a dismissive stance based largely on normative thinking about what constitutes a “proper” political organization. For better or worse, we believe that the era of large-scale, emergent, networked political organization is upon us, and, like it or not, it makes little sense to hold crowd-enabled organizations accountable to normative standards established for other kinds of organizations. In any event, these normative tangles are not to be confused with developing general grounds for empirical comparison, which we hope to have advanced in this article.

What we can offer about effectiveness and sustainability are several general observations. First, many of these crowd-enabled organizations have produced remarkable effects in short periods of time, from launching a popular constitutional assembly process in Iceland, to bringing down a regime in Egypt, to changing the discourse on inequality in the United States (Bennett & Segerberg, 2013; Castells, 2012). At the same time, the conditions that triggered the formation of such broad grassroots networks were, although very different, equally limited in opportunities for fundamental long-term political transformation. In Egypt, for example, the Muslim Brotherhood proved better organized to compete in elections, only to end up facing the wrath of the crowd when in government. Occupy did not target the 2012 U.S. elections, as many supporters had hoped. Rather, the long-term adaptive response for RQ3 above involved the formation of various local groups addressing issues ranging from supporting people facing home foreclosures (http://occupyourhomes.org) to helping victims of Hurricane Sandy (http://occupysandy.net).
Crowds surely have their limits. But in many of the political situations in which they develop, it is not clear that conventionally organized social movements would fare much better. Indeed, it may be that the crowd has certain advantages over the formal movement in terms of flexibility, regenerative capacity, and blurring the lines between core participants and bystander publics because of the ease of identification with personalized action frames (e.g., We are the 99%), and the openness of the connective action logic. These qualities may prove to be organizational advantages in an age of complex and increasingly global problems such as financial crises, climate change, food and water shortages, workers’ rights, and many other issues. Add to the proliferation of interconnected issues the incapacity or unwillingness of governments to address them in creative ways, and the traditional focus of social movements on government change is less attractive. We offer our modest theoretical framework as a point of departure for understanding and comparing organization in crowds that work in less conventional ways. The next logical step is to understand how various kinds of organization are produced through fine-grained peer production processes. A companion piece to this article by Bennett, Segerberg, and Walker (2014) explores how different peer production routines combine to produce particular types of organizational activity in the crowd.
References


