METHODS

This section provides a detailed discussion of the methods used to arrive at the conclusions in *#CommonCore: How social media is changing the politics of education*.

TWITTER DATA

Twitter (http://www.twitter.com) is a free online global social network that combines elements of blogging, text messaging and broadcasting. Users write short messages limited to 140 characters, known as *tweets*, which are delivered to everyone who has chosen to receive that user's tweets. Within each tweet, it is possible to include links to other media or to embed video, images and hashtags (a word or a phrase prefixed with the symbol #).

Twitter users can interact and communicate in different ways, and users are finding new and creative ways to get the most out of each tweet. First, they can write simple messages, called **tweets**, adding images, videos, hashtags, etc. Second, tweets can be further disseminated when recipients repost them through their timelines. This technique, called **retweeting**, refers to the verbatim forwarding of another user's tweet. A third type of messaging is a variant of tweeting and retweeting, called **mentioning**. Mentions include a reference to another Twitter user's username, also called a handle, denoted by the use of the "@" symbol. Mentions can occur anywhere within a tweet, signaling attention or referring to that particular Twitter user.

Twitter Data for the #CommonCore Project

We retrieved the data directly from the Application Programming Interface (API) in Twitter based on tweets associated with #commoncore for a period of six months from September 1, 2013, until March 4, 2014. We defined the study by the hashtag #commoncore (not case sensitive), and captured Twitter profile names as well as the tweets, retweets, and mentions using this hashtag. While #commoncore is certainly not the only hashtag related to Twitter about the Common Core (others include #cc and #ccss), it is the most prevalent tag and served as a starting point for the work. We then conducted a social network analysis of the network using Gephi[1] to identify the overall structure of this large network and then to identify subgroups and key actors that have disproportionate influence, from a social network perspective, over the information and opinions shared across the network. Our data includes messages that are public on Twitter, but not private messages between reciprocal followers.

---

1   Gephi (https://gephi.org) is a free open source software for interactive visualization, exploration and network analysis of large sizes.

Social Network Analysis

Social network analysis is grounded in the larger idea of social network theory and draws on a set of metrics to examine the pattern of connections, or ties, between individuals that create a larger social network. This network forms a social structure of relationships, which can facilitate or inhibit an individual's access to both physical and intellectual resources such as knowledge, ideas, and opinions. This structure allows for analysis at the individual, pair, small group, and overall network level and as such provides insights into patterns of interactions that are not readily visible. In the #CommonCore Project each node is an individual user (person, group, institution, etc.), and the connection between each node is the tweet, retweet, or mention/reply.

Conducting the social network analysis for the #CommonCore Project

Using the data from Twitter's API, we had to first isolate the content of the tweet itself from its associated metadata (such as a user's follower count, favorites, geolocation, etc.) and then create a file that could be read in Gephi. We then visualized the entire network including all individual actors (approximately 63,000 actors). As we were interested in only those individuals who connected to another tweeter, we narrowed the population to one giant component (a full connected network) comprised of approximately 53,000 connected actors and close to 190,000 tweets.

Determining the factions in the analysis

As we wanted to understand the inner structure and clustering of the interactions within this large connected network, we ran a community detection algorithm to identify and represent factions (a "faction" in this sense is a group with more ties within than across groups, although even those group boundaries are somewhat porous). When we ran the algorithm we found three main factions within the Common Core network.

These factions were based on the Twitter activity of the actors around #commoncore, which resulted in three distinct and overlapping groups. It is important to note, we did not "pre-assign" these factions *a priori* based on attributes of the individuals; rather, we let their interaction activity on Twitter determine the structural group to which they belonged. In other words, the content of the tweets did not influence which faction an actor was assigned to—it was based solely on an actor's ties in the network. It is also important to note that the factions are porous, meaning that the determination of an actor's "belonging" to a group is based on his or her interaction activity (meaning tweets, retweets, and mentions) with others. As such, the boundaries and membership are not hard and fast, but are rather general indicators of faction membership. We then used that data as the starting point to identify actors and then examine the ideas and beliefs of actors within factions (see section on coding of tweets).

Determining who were the key actors in the network

In order to better understand whether or not there were actors that were more active in the social network, we ran measures on each actor in order to find out which individual had relatively more incoming and outgoing ties. Having greater centrality in a network suggests an individual actor has disproportionate influence over the exchanges in that network and, as such, that his or her opinion carries more "weight." Our results suggested a number of influential actors of different types.

There are three distinct types of actors, which we call *transmitters*, *transceivers*, and *transcenders*. **Transmitters** are individuals who send out a large number of tweets using the hashtag #commoncore. Social network researchers call the activity of transmitters *outdegree*, which is a measure of the number of tweets a transmitter sends. Outdegree is *not* related to the number of followers a transmitter has, but is strictly a measure of how many tweets an individual posts using #commoncore.

Transceivers are a different kind of elite actor, those who have what social network researchers call high *indegree*. In our analyses, indegree is the combination of the number of times a person's #commoncore messages are retweeted, coupled with the number of times they are mentioned in others' tweets about #commoncore. Mentions are signifiers of importance in the #commoncore conversation.

We also identified **transcenders**, who have both high outdegree *as well as* high indegree.

Determining the smaller communities of actors

After we identified the factions and key actors in the network, we wanted to peer more deeply into the structure of the network. In order to do that in Gephi we filtered out all other actors to reach 1% of individuals with the greatest outdegree and indegree activity above the average actor in the network. We then filtered to the top .25% of the outdegree and indegree network to reveal the most highly active individuals who were over 2 standard deviations above the mean in their Twitter activity. As the data are publicly available we were then able to specifically identify the core actors and factions and conduct further analysis described in the coding section below.

Coding the Tweets

To take a closer look at the content of the tweets of the top transmitters and transceivers, we drew a random sample of 4,500 tweets (12%) of the tweets from the elite transmitters/transceivers combined and coded them in a variety of ways, including for content, political references, and choice of phrasing. A random sampling approach ensured that the resulting findings were representative of the elite transmitters and transceivers in the network.

Two undergraduate students from the University of Pennsylvania were employed in the summer of 2014 to code the tweets. The coders worked with researchers from the

Consortium for Policy Research in Education (CPRE) to develop the codes and then applied them independently to a random sample of tweets. We then met together to compare coding judgments and iterated this process, both refining the codes and discussing the responses until we gained 80% agreement between the two raters before we proceeded to code the tweets for the study.

Below are the the coding frameworks for the analyses conducted in Act 3 - The Chatter, as well as the samples used to produce the results.

Content of Common Core Tweets	Evaluates the overall type of content in the tweet. Emphasis is on the content of the tweet only, not based on author or links within tweet.	
1	CC Informational Tweets	Author provides "statements of fact," directs audience to resource, or provides an account. Since it is intended purpose, verification of fact is not an issue. This code focuses on information regarding specifically to the Common Core and its related aspects.
2	Opinions Supporting the CC	Author provides a point of view or personal commentary that is supportive of the CC. The word "should" may be an indicator (unless tweeter is quoting someone else).
3	Opinions Opposing the CC	Author provides a point of view or personal commentary that is disapproving of the CC. The word "should" may be an indicator (unless tweeter is quoting someone else).
4	Other	General catchall for those tweets that do not fall into any of the previous category. Self-promotion and rhetorical questions not focusing on CC resources or information would fall in this category.
Sample coded: Random sample of 4,500 tweets from members of elite transmitter/transceiver networks.		

Education Topics	Examines topic of tweet beyond the standards, but related to educational topics. (Note, these tweets could contain multiple references, so their total in the results may not exactly match the proportion of the sample analyzed).	
1	Teacher Evaluation	Any reference to teacher evaluation, or merit pay.
2	Testing/ Accountability	Any reference to standardized testing. May be a generic reference or to a particular testing package or regime, including PARCC and Smarter Balance, the two tests coming out for the CCSS—different states may have different names for their tests (e.g., PA is PSSA; TIMSS and PISA is for math and science).
3	Curriculum/ Textbook	Any reference to curriculum in general or specific curriculum or text-books, or topics covered by the curriculum.
4	Parents	Any reference specifically mentioning parents (e.g., moms, dads, parents) in the tweet. This also includes possessive 2nd person pronouns (e.g., YOUR children). However, this excludes tweets that only mention children/students broadly.
5	Math	Any reference to math, mathematics or any "STEM" references. Some other math words/phrases (use context to determine if it is a math term): learning progression, coherence, rigor, focus, TIMSS, PISA.
6	ELA	References to writing/listening/speech as they pertain to class activities. Some other ELA words/phrases (use context to determine if it is a ELA term): complex text/text complexity, text dependent, tier 2 words, academic vocabulary, informational text, figurative language. Includes references to writing, reading, biography, literacy, informational text.
7	Science	References to science. This may also include any "STEM" references, TIMSS, PISA.
8	Social Studies	References to social studies, history, government, or economics as content areas taught in schools. This excludes historical or government references not related to content areas.
0	None	No educational topic.
Sample coded: Random sample of 4,500 tweets from members of elite transmitter/transceiver networks.		

Political/ Policy Topics	Examines political or policy issues in the tweets. (Note, these tweets could contain multiple references, so their total in the results may not exactly match the proportion of the sample analyzed).	
1	Obama	Tweet references Obama, Barack, Barack Obama, the President, or POTUS.
2	Duncan	Tweet references Arne, Arne Duncan, Duncan, or the secretary of education.
3	Federal Role in Ed	Tweet references the governmental role in education, including terms like Govt, governmental, federal.
4	RTTT	Tweet references the Race to the Top Intitiative, may include the acronyms RTT or RTTT.
5	Gates	Tweet references the Gates Foundation, Gates, or Bill Gates.
6	Pearson	Tweet references Pearson or Pearson Publishing.
7	Data Privacy	Tweet references data, data privacy, data mining, or particular data privacy concerns like Inbloom.
Sample coded: Random sample of 4,500 tweets from members of elite transmitter/transceiver networks.		

Political Reference	Examines the topic addressed and/or person addressed in the tweet, as related to government policy (e.g., elections, laws, rulings), political theory, and/or political figures.	
1	Reference to Education Topics	Tweet makes reference to education activity that is directly connected to education but is not political in nature. (i.e., reference to testing is inside this category, but a reference to the politics of testing is not).
2	Reference to Political/Policy Issues	Tweet makes note of political figures and/or government policies that are connected to education. Political theories or ideologies are considered part of this category.
3	None	No references made to government policy or political figures.
Sample coded: Random sample of 4,500 tweets from members of elite transmitter/transceiver networks.		

Level of System	Examines the highest governmental level referenced in education topic or politics/policy-related tweets.	
1	International	References something outside of the United States
2	National	References the national level
3	State	References the state level
4	Local	References district or school levels
5	Unspecified	Level cannot be determined
Sample coded: 930 tweets that were coded as either education topic or politics/policy-related issue.		

Analysis of Policyspeak vs. Politicalspeak

As we examined the data, we noticed that some of the tweets used neutral language and focused on policy aspects of the Common Core (policyspeak), while others were much more emotionally charged (politicalspeak). We wondered if these types of messages were associated with the different factions in the structural networks.

We sought to explore the relationship between the structural faction and this aspect of the language of the tweets by coding the data for policyspeak or politicalspeak. To conduct this analysis, we drew a stratified random sample of tweets from the 930 that referenced education topics or politics/policy issues. We sought to avoid sample bias by stratifying equally by faction. Because this sample was heavily weighted toward tweets from the faction of actors outside of education (yellow), we took the lowest represented group (the blue faction, which contributed 168 of 930 tweets) and drew equivalent random samples for the green and yellow factions. This produced a total sample of 504 tweets. We then coded the 504 tweets according to the following rubric:

Type of Speak	Examines the nature of the tweet in terms of rejection or acceptance of the common core at different levels.	
1	Policy Speak	Tweet seeks acceptance or rejection of the Common Core by focusing on Common Core as policy (e.g., implementation, proposed outcomes, evaluation, precedent). Tweets can exhibit this type of speak through policy jargon, referencing data/precedent, similarity/differences to other policies. Language tends not to be inflammatory or loaded. Lack of "call to action." Does not necessarily call for outright rejection or acceptance of CC, but focuses on refinement/alternatives.
2	Political Speak	Tweet seeks broader acceptance or rejection of the Common Core (and/or related supporters or opponents) as an entire idea/movement. Tweets can exhibit this type of speak by a "call-to-action," symbolism, loaded (emotional) language, and/or inflammatory language in addition to making broader statements about the CC and supporters/opponents.
3	Undetermined	Tweet may be unclear in meaning, advertisements, or providing information that does not relate to acceptance or rejection of the CC.
Sample coded: Stratified random sample of 504 tweets; stratified by faction so as not to *a priori* bias results.		

To conduct statistical analyses of the differences between factions, we used the resulting coded data and conducted separate analyses of variance for each type of speak by faction, with a post hoc test of differences between groups.

INTERVIEWS

The nine interviews in Act 4 – Motivations, were conducted at the end of 2014 and beginning of 2015. The interviews were conducted via telephone with individuals and organizations in the elite transmitter/transceiver networks. We originally reached out to 12 people/groups that were represented in the three factions in the social network. To contact someone, Supovitz followed the user on Twitter in the hopes they would reciprocate so that he could send them a private message inviting an interview. In other cases, he searched for their contact information on the Internet. Three of the sample either never responded to interview requests or declined to be interviewed. Due to audio problems in the recorded interviews of two people, we did not produce podcast for them. We make no claims as to the representativeness of the final sample, but their interviews both enriched the picture of different aspects of the Common Core debate and/or illustrated different themes that had surfaced in other aspects of the data.

Interview Protocol

The interviews were semi-structured, whereby a set sequence of questions was followed while also allowing for latitude to probe and follow up on issues raised by the respondent. The following is the interview protocol:

Hi, my name is Jon Supovitz and I am a researcher at the Consortium for Policy Research in Education at the University of Pennsylvania. I've been studying the Common Core debate on Twitter and notice you are a very prolific voice in that discussion and that's why I reached out to you.

I'd like to interview you as part of a research project that will produce interviews for a website examining the Common Core debate on twitter. May I have your permission to audiorecord the conversation? If there are things that you prefer I not record, I will be happy to shut off the recorder at any time during the conversation, just let me know.

1. So tell me a little about yourself and your background.

2. I see on Twitter that you are involved in a variety of issues, so what got you so interested in the Common Core?

3. Do you recall any particular thing that catalyzed your interest?

4. How would you describe the Common Core to someone who was unfamiliar with the topic?

5. What are some of the other issues you are involved with? Where does Common Core rank in a list of the issues you engage in on Twitter (approximately)?

6. Why do you think the Common Core is such a contentious topic?

7. Has your position on Common Core changed since you engaged in the conversation on Twitter? If so, how?

Now let me focus a bit on social media and Twitter:

8. How frequently are you on Twitter?

9. Do you have any particular strategies that you use to be so central to the Common Core conversation on Twitter (specific hashtags, links)?

10. Do you use any other mediums beside Twitter? Facebook? Blogs, etc? Why did you choose Twitter to talk about the common core? (Focus: Twitter as medium)

11. In what ways do you think Twitter is changing peoples' opinions about the Common Core? How do you know?

12. Do you think Twitter is changing the political conversation in the country?

Thanks for your time. I will reach back out to you to share the product of this interview with you before posting it on our website to give you the opportunity to react.

Interview Analysis

Once the interviews were completed, Supovitz listened to each recording multiple times and selected multiple excerpts that (a) described the individual or group's interest in the Common Core; (b) depicted the motivation for their support/opposition, and; (c) described their use and views of social media and Twitter as an interactive communication platform. These became the rough cuts for the interviews. He then recorded the questions and comments, and these were inter-spliced with the interview segments.

After the interview podcasts were completed, they were put on Dropbox for each respondent to hear, after which Supovitz and the interviewee communicated about and resolved any questions that arose to the respondents' satisfaction.