The Structure and Characteristics of #PhDChat, an Emergent Online Social Network

Kasey C. Ford
Texas State University
kaseycford@gmail.com

George Veletsianos
Royal Roads University
veletsianos@gmail.com

Paul Resta
University of Texas at Austin
resta@mail.utexas.edu

Abstract: #PhDChat is an online network of individuals that has its roots to a group of UK doctoral students who began using Twitter in 2010 to hold discussions. Since then, the network around #PhDchat has evolved and grown. In this study, we examine this network using a mixed methods analysis of the tweets that were labeled with the hashtag over a one-month period. Our goal is to understand the structure and characteristics of this network, to draw conclusions about who belongs to this network, and to explore what the network achieves for the users and as an entity of its own. We find that #PhDchat is a legitimate organizational structure situated around a core group of users that share resources, offer advice, and provide social and emotional support to each other. Core users are involved in other online networks related to higher education that use similar hashtags to congregate. #PhDchat demonstrates that (a) the network is in a continuous state of emergence and change, and (b) disparate users can come together with little central authority in order to create their own communal space.

Keywords: Online networks, social media, online participation, Twitter, social networks, #PhDchat, hashtag, higher education, emergent online communities, networked participatory scholarship

Introduction

Though all Social Networking Sites (SNS) allow for rapid collaboration and exchange, perhaps none is as effective at facilitating bursts of dialogue as Twitter, the microblogging platform that allows users to publish short strings of text and curate their own feeds made up of updates from other users. Twitter has become a useful tool for individuals and organizations, as it provides a participatory space through which participants can self-organize, converse, and distribute their messages to reach user networks (Java, Song, Finin, & Tseng, 2007).

In this paper, we examine the network formed around the #PhDChat hashtag. #PhDChat was originally developed as a way for UK-based doctoral students to hold weekly discussions. Nowadays, the hashtag is added to hundreds of tweets per day and the network has morphed into a vibrant participatory space used by numerous individuals (doctoral students and otherwise). The real-time weekly discussions that generated the moniker continue each Wednesday evening and now include participants from around the world.

We chose to study the #PhDchat hashtag and network for a number of reasons, including:

- We are interested in examining learning, teaching, and knowledge creation/dissemination practices in networks, and #PhDchat represents a naturalistic setting in which these practices occur.
- #PhDchat has formed organically and appears to have little in the way of a central structure.
- The characteristics of this network and, in turn, the characteristics that it has in common with other emergent online networks will generate insights into how people are using the Internet to create their own learning opportunities and form social support networks.
- Analyzing this network will contribute to our understanding of how and how effectively knowledge exchange and dissemination are occurring online.

The research question we will answer in this paper is the following: What is the structure and characteristics of the network that has formed around the #PhDchat hashtag on Twitter? To answer this question we analyze the discourse that was labeled with #PhDchat over a one-month period, using a mixed methods approach. We first review literature relevant to the topic. Next we describe our data collection and data analysis methods. Finally, we discuss our findings and present implications for future work.

Review of Relevant Literature

Social Networking Services (SNS) have had an effect on the way people consume news (Glynn, Huge, & Hoffman, 2012), engage in the political process (Gil de Zúñiga, Nakwon, & Valenzuela, 2012), and create social circles (Thompson, 2008). Boyd & Ellison (2007, p. 211) define SNS as "web-based services that allow individuals to (1)

construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system". The contributions of these technologies to learning are also promising. Jenkins, Clinton, Purushotma, Robison, and Weigel (2006, p. 3) for example, argue that the participatory cultures forming around social media promise "opportunities for peer-to-peer learning, a changed attitude toward intellectual property, the diversification of cultural expression, the development of skills valued in the modern workplace, and a more empowered conception of citizenship". Growing interest in SNS has also cultivated fertile ground for educational research (Greenhow, Robelia, & Hughes, 2009). For example, researchers have examined the implementation of Twitter in classrooms (Young, 2010) and in informal learning contexts (Aspden & Thorpe, 2009).

Researchers have argued that a set of skills and proficiencies are necessary if social media are to provide participants with more and better opportunities to learn (Jenkins et al., 2006; Rheingold, 2010). For instance, Rheingold (2010) argues that shifting between multitasking and focused attention is a skill that has become essential to learning effectively in today's digital environments. Blankenship (2011) suggests that social media can encourage educators to think more creatively about teaching and learning, but effective integration into classrooms depends not only on taking advantage of the opportunities provided by the tools, but also ensuring user proficiency with social media.

The research literature on social media use in education is broad, largely because these technologies have been used for multiple purposes (e.g., instructional vs. research uses), within different contexts (e.g., formal vs. informal learning), and by different actors (e.g., individual vs. institutional use). For example, researchers have examined local SNS created to help transition incoming freshmen into their college careers (DeAndrea, Ellison, LaRose, Steinfield, & Flore, 2012), investigated the sharing of school-related knowledge on online social networks (Wodzicki, Schwämmiein, & Moskaliuk, 2012), explored the use of online social networks by faculty (Kaya, 2010), and studied the integration of social networking environments in traditional higher education settings (Veletsianos, Kimmons, & French, 2013).

One social media technology that has attracted significant attention in the research literature is Twitter. At the time of writing, Twitter was used by approximately 16% of Internet users (Duggan & Brenner, 2013) and, like other SNS, found its way into higher education settings. The tool has been described as being valuable for both instructional (Dunlap & Lowenthal, 2009) and scholarly (Veletsianos, 2012) purposes. Researchers have argued that it enables effective peer-to-peer communication (Kassens-Noor, 2012), cultivates ongoing dialogue (Lalonde, 2011), allows opportunities for sharing, eflecting, and discussing (Ebner, Lienhardt, Rohs, & Meyer, 2010), and fosters active learning (Junco, Heiberger, & Loken, 2011). Furthermore, Twitter has been identified as a professional development tool (Gerstein, 2011), especially amongst teachers (Ferriter, 2010; Forte, Humphreys, & Park, 2012; Holmes, Preston, Shaw, Buchanan, 2013). For example, Twitter-using educators frequently indicate that they use the platform to

create professional ties and share resources (e.g., Forte et al., 2012). Similar results have been reported by Veletsianos (2012) who studied scholars' tweets and found that individuals ask for and provide resources, assistance, and advice to students and colleagues alike.

The integration of Twitter in teaching and learning contexts is not without challenges. For example, Kassens-Noor (2012) suggests that the tool does not provide significant opportunities for self-reflection and Petrilli (2011) notes that a SNS may simply function as a soapbox. In recognition of these issues, Lin, Hoffman, and Borengasser (2013) highlight the need for proper scaffolding, allowances for privacy, and explicitly-stated purposes before using Twitter in a course. Finally, Veletsianos and Kimmons (2012) argue that online social networks may mirror issues of power and class and even though they may be promoted as tools for collaboration and dialogue, they may not necessarily foster equality and democratization.

Increasingly, online social networks, including Twitter, appear to become places used by individuals in order to collaborate (), share intimate details of their life (Thompson, 2008), and connect with others (). In this way, online social networks become places of gathering (Veletsianos, 2013) and places that create opportunities for creating, cultivating, and sustaining relationships. However, the literature does not provide a clear understanding of what these online places look like, especially in the context of social networking sites and other platforms with ephemeral communication mechanisms. The research presented in this paper reduces the lack of knowledge in the area using social network analysis, which is a method suggested by recent research as helpful to consider in determining the form of online environments (Gruzd & Haythornthwaite, 2011).

Further, emerging evidence suggests that the use of hashtags (a common Twitter practice allowing users a means to group and retrieve messages around a common topic) can foster building and maintaining of relationships (Gruzd & Haythornthwaite, 2011; Reed, 2013). Perhaps less clear are the ways in which emergence, or the process through which participants self-organize through the use of a hashtag, impact the substance of an online community. While research on networked learning has discovered that learners curate their own personal learning networks (e.g. Couros, 2010), there is little research describing what happens when learners organize themselves spontaneously (Dron & Anderson, 2009). Yet, the literature on online learning communities broadly has a wide research base that we can draw upon to generate insights on what gatherings around SNS might look like. Riel & Polin (2004) for example, categorized learning communities into three types:

1. Task-based - members are assigned according to task features; clearly defined project or problem with a start and finish.
2. Practice-based - arises around a profession or discipline; learning is the result of ongoing practice.
3. Knowledge-based - participation Online arises out of relevant expertise or common interest; knowledge base evolves.

Each one of these communities has different needs and faces different challenges.

Importantly, since web services - in conjunction with avenues for discovery like search engines, social bookmarking sites, and advertisements - allow users to come together from disparate geographical locations and interest groups in order to form communities, the resulting demographics are often more diverse than local or institutionally organized groups, leading to organizational systems that are "complex, fractal and turbulent" (Doll, 2009, p. 164). This complexity is important to highlight, as it also appears to be present in open learning and scholarship environments. There is still much to understand about emergent online learning spaces, how they are organized, and how they foster understanding and create bonds between disparate groups of people. To contribute to this understanding, we examine #PhDChat in order to identify the attributes of this self-organizing network, identify its major characteristics, and identify the ways it is used by its members.

Context

Twitter

This study occurs in the context of Twitter. Twitter is both a social networking site and a microblogging platform (Veletsianos, 2012), as it allows users to (a) follow each other, and (b) post text updates (called tweets). A tweet can consist of a combination of 140 ASCII characters and once submitted, the tweet will either be posted publicly and aggregated into the timeline of the user's "followers" or become available to those followers whom the user has given permission to read their tweets if the user has set their profile to private. Twitter users often use the service to chat, converse, share information/URLs, and report news (Java et al., 2007). Casual observers are often surprised to discover that extensive conversations occur on this platform ().

SNS engender their own communication styles and social interactions (Herring, 2008), and one of Twitter's most common practices is the use of the *hashtag,* which is a simple "#" symbol followed by a word or phrase (e.g., #fun, #StateOfTheUnion, #elections, #education). This practice allows users to tag a message (e.g., "I am enjoying meeting colleagues at the #aect2014 conference"). This form of social tagging provides a means to group and retrieve messages around a common topic. For instance, users who tweeted about watching the World Cup final might include the hashtag #WCFinal in their tweets and those who were interested in following public reaction to the event could conduct a search simply by clicking on a hyperlinked hashtag. This practice has allowed users to instantly and autonomously form networks around shared interests (Parker, 2011) such as entertainment, events, sports, political causes, jokes, and legislation.

Twitter also allows users to republish the tweets of others as retweets. The text "RT" is automatically added to tweets when a user selects the retweet button on the tweet they would like to share with their followers. A modified tweet, or MT, is a tweet that has been marginally edited by a user (e.g., by adding a hashtag or a comment to the message). Users frequently indicate that a retweet has been modified by replacing RT with MT.

#PhDChat

#PhDChat is the hashtag and network that we examine in this paper. #PhDchat has been mentioned as an example of a community that other academic interest groups might emulate (Coiffait, Bartlett, Houghton, & Condie, in press). Unlike other hashtags, the origins of #PhDchat are unambiguous and the history of the community is well-documented. According to the wiki for the community, the hashtag began when a group of UK doctoral students started using it in 2010 as a way to hold discussions over Twitter (Thackray, n.d.). A planned discussion about a specified topic was still occurring each Wednesday at 7:30 GMT each week at the time this paper was written. Indeed, the discussion that the group holds each Wednesday differentiates it from other networks and hashtag-using groups. #PhDchat has evolved since its original inception and the hashtag has been used by individuals outside of the core group of original participants for regular communication outside of the planned discussions. As a result, #PhDchat consistently appears in tweets outside the weekly chat.

Research Questions

We answer the following research question: What are the structure and characteristics of the network that has formed around the #PhDchat hashtag on Twitter?

Methods

This mixed methods study uses tweets that included the #PhDChat hashtag in order to identify the characteristics of an emergent online network of Twitter users. The study relies largely on quantitative data in the form of social network analysis and statistics in order to draw conclusions about the community being studied. Nonetheless, online networks are inherently social and can rarely be wholly quantified. Where appropriate, we elected to present and examine individual tweets for meaning and make observations about the interactions that took place between users in order to provide a more holistic picture of the network.

Data Collection

All of the public tweets that contained the text string "#PhDchat" were collected during 39 days in 2013. This archiving method focuses on individuals who used the #PhDchat hashtag and excludes individuals who may have engaged with the network but in a manner that did not include use of the hashtag (e.g., lurkers). The specific time period used was chosen because it fell within the scope of a traditional semester, but did not coincide with the beginning of, the end of, or a break in classes. We estimated that tweets during the beginnings, ends, or breaks were special times that could result in unique levels of participation, and even though these unique time periods present interesting opportunities for research endeavors, we wanted to avoid the uniqueness of specific time periods having an impact on our results. The duration of about one month was selected because a preliminary data sample collected to pilot the study revealed

that a one-month period provided a large but manageable amount of tweets for analysis. All tweets collected were included in the analysis and no tweets were removed from the data pool. The raw text of the tweets was collected through the third party web service that was available at the time (http://www.tweetarchivist.com) that enabled us to retrieve and archive tweets.

The tweets collected as data in this study are available publically through Twitter or any other application that utilizes the Twitter Application Programming Interface (API) for data retrieval. We sought and obtained a non-human subjects research waiver determination from our institution's Internal Review Board, as the tweets collected were publicly available, posted at the user's own volition, and the study posed no risk to them in addition to the risk they assumed upon agreeing to Twitter's terms of service and choosing to publish their tweets publically. Nevertheless, we took additional steps to further minimize potential risks to users. In particular, in our archived data set we obscured Twitter account information, removed identifiers, obscured URLs that may have given information that revealed the identity of users, and modified tweets used in this paper to avoid identification if one were to search for them using a search engine. Even though the tweets we use in this paper to illustrate the results differ from the original, we compared them to the original to ensure that the revised versions maintained the original intent.

Data Analysis

Tweets were downloaded in plain text, comma delimited format for analysis. Usernames were replaced with randomly assigned identifiers consisting of the word USER and a four-digit number. Geographic location was discarded. After the data were cleared of all identifying information, a spreadsheet application was used to calculate the basic network statistics including hashtag frequency, languages used, tweet source, number of users mentioned, number of users who tweeted, and number of users who participated in similar groups. Tweet dates and timestamps were separated into a different worksheet for additional coding and a histogram was created to determine the frequency of tweets during each hour of the day.

Word frequency analysis was performed on the dataset after the initial calculation of network statistics. Tweets were copied into a text file and the #PhDChat hashtag and assigned identifiers were removed from each entry. A spreadsheet operation was used to break each tweet string into individual words, and another operation was run to generate a list of all of the words found in the dataset. Once the list of unique words was compiled, prepositions, pronouns, possessive pronouns, symbols (e.g., | and @), all forms of the verb "be," transitive verbs, "RT", "MT", and conjunctions, were removed from the dataset and a count function was used to count the number of instances of each word.

The Microsoft Excel-based Node XL software was used to create network visualizations in order to enhance our understanding of the connections, relationships, and groups within #PhDchat. In each case, the default NodeXL graph options were used. Detailed

information on the algorithms used for each graph is provided in Appendix 1.

Results

Network Statistics

12,723 public tweets that contained #PhDchat were collected. These tweets were posted by 3,299 users. Between those who tweeted or were mentioned in a tweet, 4,102 users directly or indirectly participated in the PhDchat community. Hashtag-using participants included frequently-contributing community members as well as individuals who only used the tag once. The 20 most prolific users contributed about 27% and the 100 most prolific users contributed about 48% of all tweets in the dataset (Table 1). These individuals are the core users of the network and, unsurprisingly, all but one of these users also participated in the Wednesday night discussions. 2,106 users (the majority) only contributed one tweet during the study's date range. Of the 2,106 users who only contributed one tweet, some may be infrequent contributors while others may have simply retweeted a tweet that included the hashtag.

Table 1. Top 20 most prolific users

Twenty Most Prolific Users	Total Tweets	Percentage of Total Tweets
USER4052	431	3.39%
USER2749	386	3.03%
USER3895	331	2.60%
USER4643	256	2.01%
USER2206	165	1.30%
USER1287	155	1.22%
USER4876	149	1.17%
USER2092	148	1.16%
USER2898	122	0.96%
USER1872	111	0.87%
USER4290	111	0.87%
USER1221	106	0.83%
USER2309	103	0.81%
USER4196	98	0.77%
USER4733	94	0.74%
USER1369	88	0.72%
USER2368	87	0.68%
USER1173	79	0.62%
USER4458	79	0.62%
USER4331	75	0.59%
TOTAL	3174	27.06%

There were 74 and 3,944 MTs and RTs, respectively, which accounted for 31.58% of all tweets. In addition to re-sharing and echoing content in the form of RTs, users often shared URLs. A total of 2,352 unique URLs were shared in 5,105 or 40.12% of the total tweets collected. It is necessary to note that the tendency to employ URL shorteners in order to conserve characters for length-limited tweets could mean that different URLs could direct readers to the same location.

Language

The default language of the user was included in this data set. Approximately 99% of the tweets in the dataset were drafted in English, and users with alternate default languages wrote 129 tweets. The languages were Polish, Danish, Indonesian, French, German, Spanish, Portuguese, Arabic, Vietnamese, Lithuanian, Japanese, Italian, and Swedish. Of the tweets that were posted by users who used a default language other than English, 95 were composed in English and 34 were in a language other than English.

Source

The source (operating system, client, or program used by the user to publish each tweet) for a majority of the #PhDchat tweets was the Twitter website and accounted for 37.48% of all entries. Twitter for iPhone was the next most popular platform with approximately 13% of tweets posted, and TweetDeck was third most popular, contributing more than 11% of traffic. Five of the top 10 traffic sources were exclusively mobile and made up about 28% of all tweets.

Table 2. The top 10 sources of #PhDchat tweets

Source	Total Tweets	Percentage of Total Tweets
Web	4768	37.48%
Twitter for iPhone	1648	12.95%
TweetDeck	1431	11.25%
Twitter for iPad	726	5.71%
HootSuite	653	5.13%
Android	583	4.58%
Tweetbot iOS	395	3.10%
Tweet Button	287	2.26%
Buffer	252	1.98%
BlackBerry	227	1.78%

Hashtags

Though all of the tweets in the dataset contain the #PhDchat hashtag, 1,754 other unique tags were present, and these were used a total of 14,333 times. Approximately 47% of the total tags used in all tweets were #PhDchat tags. The next top 20 tags accounted for about 30.5% of all hashtag uses and about 58.5% of hashtags not including #PhDchat.
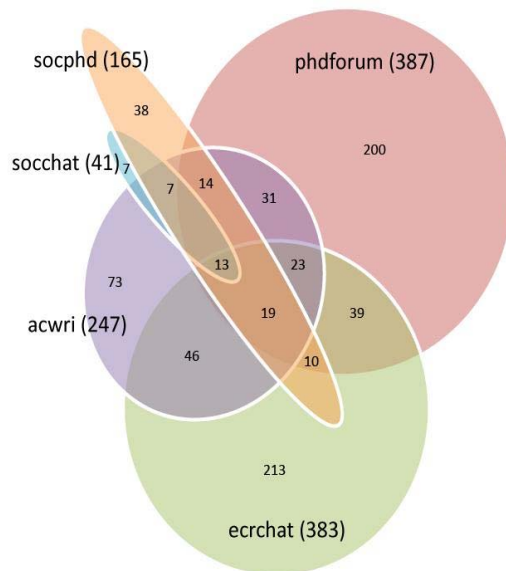
Table 3. Top 20 hashtags other than #PhDchat

Hashtag	Total Instances	Percentage of Total Non-PhDchat Instances
phdforum	1635	11.41%
phd	975	6.80%
highered	823	5.74%
ecrchat	792	5.53%
socphd	716	5.00%
acwri	663	4.63%
phdadvice	514	3.59%
dissertation	337	2.35%
academia	326	2.27%
research	274	1.91%
gradchat	215	1.50%
gradhacker	213	1.49%
socchat	166	1.16%
thesis	155	1.08%
writing	126	0.88%
edchat	119	0.83%
lovehe	118	0.82%
gradschool	86	0.60%
ecr	71	0.50%
education	70	0.49%
Total	8394	58.58%

The 20 most popular hashtags shown in Table 3 can be divided in two categories [1]: tags loosely associated with #PhDchat and tags used to highlight a topic. The tags associated with #PhDchat are often bound by organizational structures. For example, #phdforum refers to a group that connects those in higher education. The tags #socphd and #socchat are associated with #phdforum and focus on social research. #ercchat is a network very similar to #PhDchat in that it holds chats for users weekly, but is more focused on the issues of early career researchers. The #acwri community holds bi-weekly chats and is geared towards academic writing. Figure 1 shows the overlaps in the use of the five tags related to other communities (#PhDforum, #socphd, #socchat,

#ECRchat, and #acwri). Results show that 797 of individuals used these five hashtags in addition to #PhDchat in at least one tweet. While only 13 users used all five tags, there was significant overlap between users who used certain combinations of tags. For instance, 34 of 41 users who included #socchat in a tweet also used #socphd at some point. Almost half of all users that used #acwri also used #ecrchat as well.

Figure 1. Proportionate use of tags related to #PhDchat among users

socphd (165)
38
phdforum (387)
200
socchat (41) 7
7 14 31
13 23
acwri (247)
73
19 39
46 10
213
ecrchat (383)

Within the tags loosely associated with #PhDchat we include looser organizational structures like #Gradhacker (an individual and associated group using a twitter feed, blog, and hashtag to post resources for graduate students) and #phdadvice (a group without a regular forum or its own webpage used by individuals seeking the advice of their colleagues).

The second kind of tag found in the list in Table 3 represents less formal tags used to highlight a topic, such as #phd, #dissertation, #academia, and #writing, which are common words transformed into annotations by users. Users often appended the # sign before words to highlight them. Examples include:

"Can anyone suggest some good books for #PhD educational research? Any advice would be great :) #PhDchat"

"Good meeting. My advisor read my full #dissertation draft. I have my orders. Now to finish ANOTHER draft. #PhDchat"

Less frequently used hastags in the list of 1,754 tweets show signs of playful asides

(e.g., #longlivethepjs, #postdocalypse, and #overlyhonestmethods) or explanatory remarks (e.g., #nervous, #worklifebalance, #phdprobs, #revolting, and #notproductive) that many Twitter users embrace in order to add meaning to their character-limited entries.

Engagement over time

Tweets retrieved included a timestamp, indicating the time that each tweet was posted. An analysis of timestamps revealed that tweets were published steadily, but would slowly rise during traditional work hours (9:00 AM to 5:00 PM GMT) on weekdays. Saturday and Sunday yielded lower total tweet counts overall with less of an upward trend during work hours. Wednesdays revealed a similar pattern, but because they were the day during which live chats were scheduled, they drew the greatest numbers of tweets (3,409 of 12,723), and included a sharp rise in number of tweets at the beginning of the live chat session (figure 2).

Figure 2. Total tweets per day, divided in four six-hour ranges
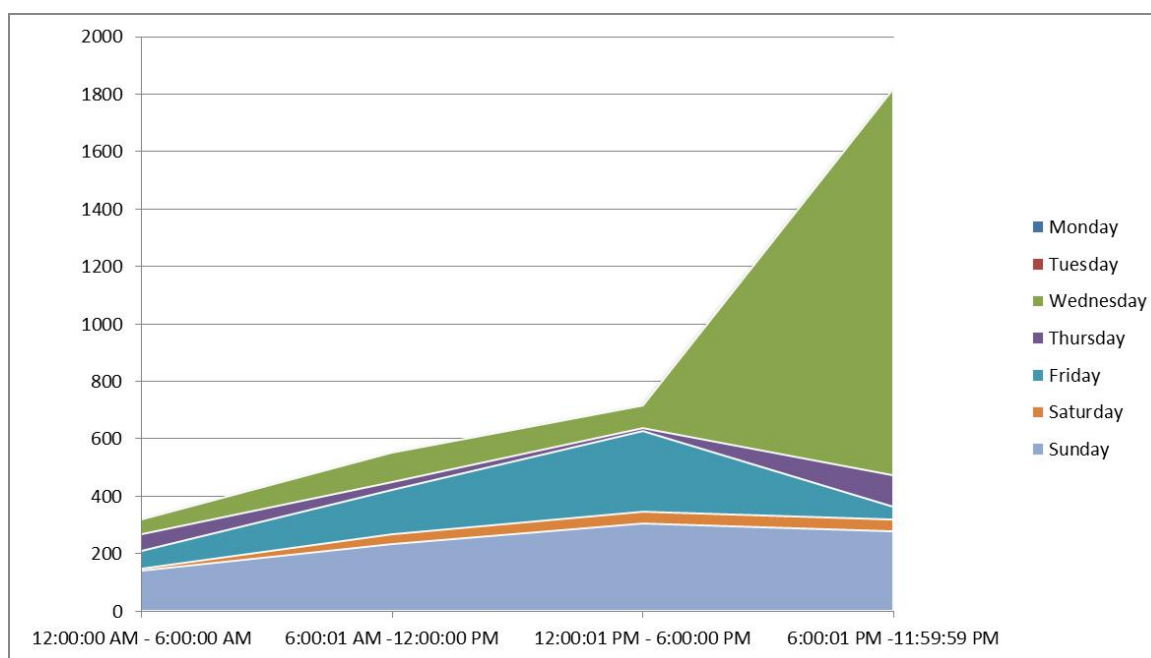
A comparison of the frequency of timestamps regardless of day reveals that entries spike during the #PhDchat discussions (8:00 and 9:00 PM GMT). This comparison also reveals that tweets in the late evening were more prevalent than those early in the morning. For example, there were more tweets published at 12:00AM, 1:00AM, or 2:00 AM than at 8:00 AM on any given day.

Figure 3. The number of tweets per hour visualized
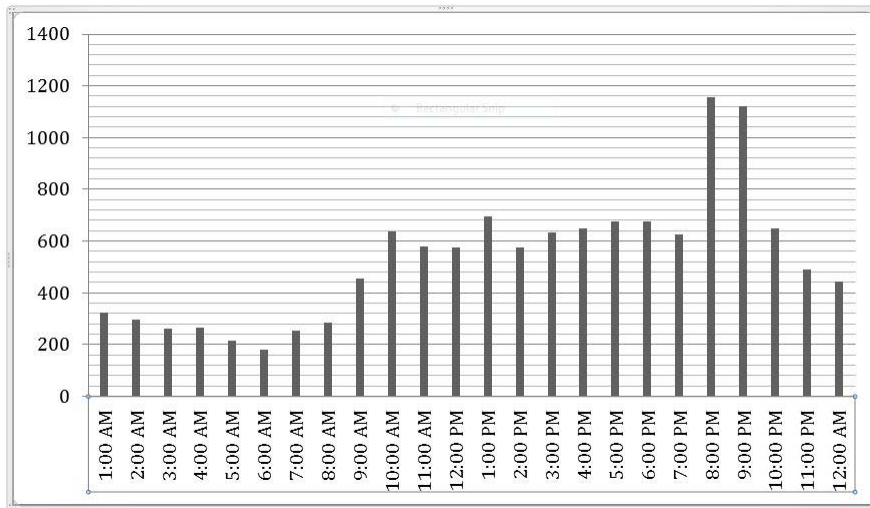
Since it is likely that many of the users that are tweeting from 7:30 to 8:30 PM GMT on any Wednesday are likely to be participating in the group discussion, we felt it was important to isolate this information. The number of tweets during Wednesday discussions ranged from 149 to 250 (average of 178) and the number of users participating ranged from 32 to 41 (average of 37).

Word Frequency Analysis

Word frequency analysis of the text contained within the tweets revealed that more than 14,000 unique words were used. #phdcat and a number of other hashtags were amongst the most frequently used words. Since conversation surrounding the process of pursuing a PhD was a common topic of discussion, frequently used words were associated with these topics (e.g., research, academic, writing, thesis [2]). Words related to the pursuit of a higher education degree, such as "data" and "reading," were also numerous. "Methods", "analysis", "article", and "conference" occurred less frequently, but still appeared in the text hundreds of times. Terms such as "tweet", "Twitter", "post", and "blog" were also common as the lexicon of the medium. The list of words from this analysis also contained references to science, literature, engineering, and the social sciences, suggesting that this network is composed of individuals from multiple disciplines.

Figure 4. A word cloud generated by the text of the tweets containing #PhDchat

Social Network Analysis

The #PhDchat network contained 11,184 user mentions in 7,798 (out of 12,723 total) tweets. To better understand the structure of this network we used social network analysis to understand the relationship between participants. This analysis is portrayed in figures 5, 6, and 7. In these figures individuals are represented as nodes, and interactions between individuals are represented as lines (ties) between nodes. The ties represent either a 1-way interaction or a 2-way interaction. We did not include the direction of the interaction in the visualizations because its inclusion impeded clarity and did not provide additional helpful information that was not already provided by the analysis that precedes this section. The coloring of the nodes is insignificant and only serves to make the visuals more convenient to scan.

Figure 5 shows the #PhDChat network divided in clusters. Users with frequent or exclusive ties, represented in this study as replies and mentions, are clustered together. Thus, each cluster represents users that are most closely associated to one another based on their frequency of interactions. The small clusters at the top right-hand side of the figure represent individuals who interacted with a small number of other individuals in the network (one to three usually). These clusters often tie back to the major clusters shown in the left-hand side and bottom half of the image. The connection to the major clusters is often the result of a user re-tweeting an account with a large following and then engaging in a brief interaction with followers of that account. This activity pulls in users who are otherwise not active in #PhDChat and thus appear in their own separate

clusters in the visualization. Figure 5 also shows that the network:

- consists of several major groups of users, tightly clustered together through replies and mentions;
- contains several smaller isolated or loosely connected groups; and
- has about a dozen large clusters of users and many smaller, less densely-connected ones.

In addition, (a) the majority of the top 100 most prolific users appear in the largest and most centrally connected group, and (b) most groups have significant ties back to the largest cluster; in fact, the dense group of users is the only one tied to the smaller groups.

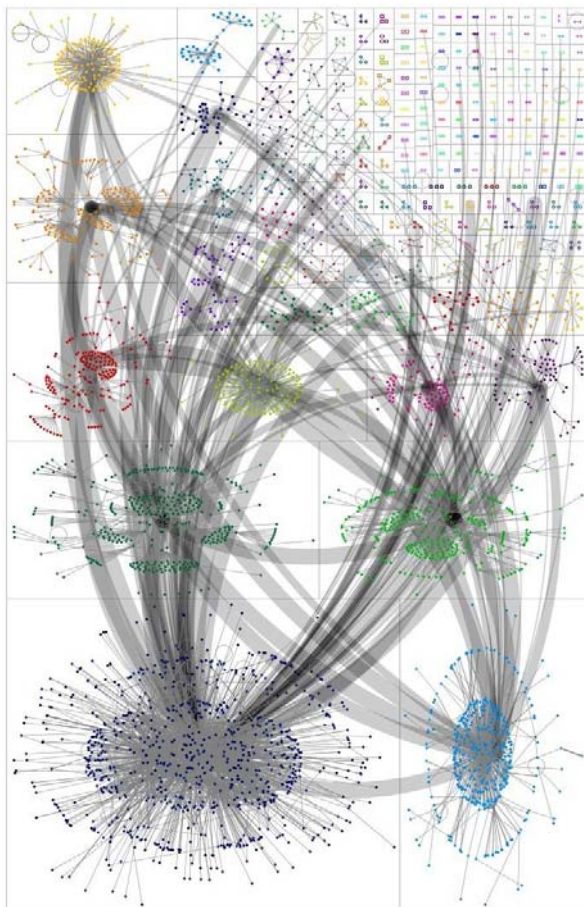Figure 5. A visualization of all mentions in #PhDchat with users grouped into clusters*[i]*

Figure 6 shows the network without the clusters/groups. This image shows that some users fall outside of the purview of the core group of participants and, during the period of data collection, had no interactions with the larger, more densely connected community. By removing the users who contributed or were mentioned in only one tweet (figure 7) we see that the peripheral nodes mostly disappear, leaving a more

tightly associated group of active individuals.

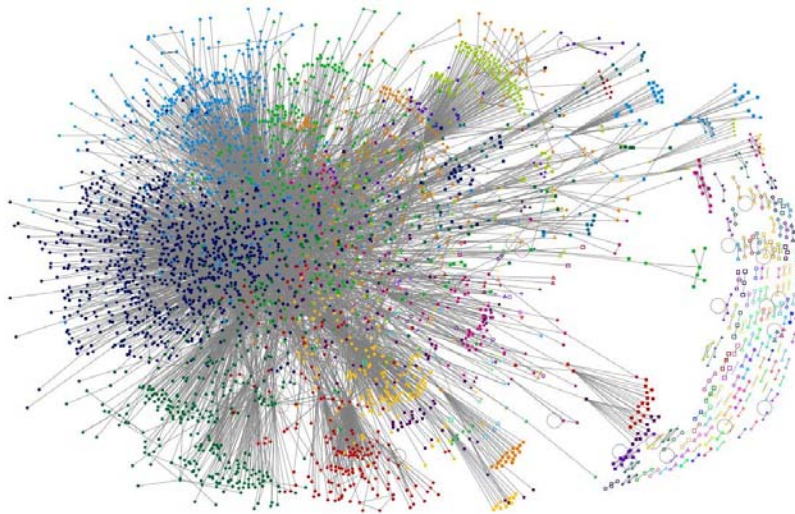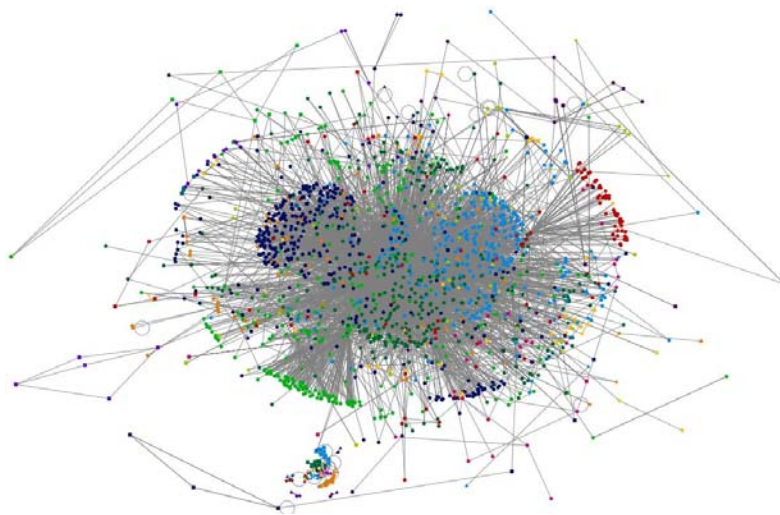Figure 6. Users who mentioned or were mentioned in a tweet containing the #PhDchat hashtag[ii]

Figure 7. All mentioning and mentioned users with more than one interaction[iii]

Discussion

The data collected for this study reveal a significant amount of information about the #PhDchat network. Participation patterns suggest that even though users tweet throughout the day, many of them also keep late hours with tweets tapering off into the early morning hours. This time period suggests that many of the users operate in or close to Greenwich Mean Time, and therefore are located in Western Europe and Africa.

However, the significant use of the term dissertation suggests that there are also a large number of North American users. At least one quarter of them also appear to use a mobile device. Though the common words in the dataset suggest that users were engaged with writing dissertations/theses, a close reading of tweets suggests that the network engages with numerous aspects of the doctoral experience including sharing the trials, joys, and day-to-day happenings of pursuing a higher education degree. For example, participants expressed their frustrations (e.g., "I have written 8 words over the past two hours. EIGHT. #epicfail #PhDchat #thesis"), asked for advice (e.g., "I need to put together a teaching philosophy and teaching pack. Any suggestions on resources or SAMPLES? #PhDchat #PostDoc #PhDForum"), shared resources (e.g., "What to expect from your first teaching assessments: popular article {URL} #highered #PhDchat"), and reflected on their work (e.g., "#PhDchat the PhD was trying, esp in the last few years, great to have my passion back").

While our analysis of word frequencies highlights prevalent topics of discussion, it does not capture the tone of communication. In addition to the numerous tweets sharing resources, many of the messages were supportive and conversational. Inquiries from individuals were frequently answered with numerous responses and the formal discussions on Wednesday tended to spur rapid and detailed exchanges. At times, #PhDchat participants also tried to inspire and encourage others (e.g., "Let's write more than a tweet today folks! #writing #highered #PhDchat #phdstudent").

The top hashtags in the dataset, and their frequencies, indicate connections between #PhDchat and similar groups, and introduce a number of questions. For example, does use of one of these hashtags make one more likely to use the other ones? The five most frequently used hashtags (#phdforum, #socphd, #socchat, #ecrchat, and #acwri) represent interest groups similar to #PhDchat, and their frequency within the dataset is unsurprising. It is also unsurprising that the two closely related tags #socphd and #socchat indicate a large overlap in their user base. Another question that arises is: What motivates users to use multiple hashtags? Pragmatic reasons might be behind the practice of using multiple hashtags, as this allows users to bring their tweet to the attention of different groups of people monitoring different hashtags. This practice may not work well with closely-associated hashtags (e.g., #socphd and #socchat), but may work well with loosely-associated hashtags (e.g., #PhDchat and #lovehe). Adding multiple hashtags to a message might also be a network-building strategy or a strategy to broker information between communities that might not otherwise interact much with each other.

What sort of organizational structure describes the individuals using the #PhDchat hashtag? Do these individuals belong to a community, a network, or an interest-driven group? Using a hashtag does not necessarily create a community out of otherwise unrelated individuals. For example, individuals who tweet about the Olympic opening ceremonies and use #OlympicsOpeningCeremonies as a hashtag may be part of a loosely associated group as they read comments, respond, and remix the content of others, but this group of people are not necessarily a community that shares a sense of belonging to the group (McInnerney & Roberts, 2004). It is unlikely that the people

tweeting in the hypothetical situation described above, though they are gathering within the social structure of Twitter, feel a sense of belonging because they contributed to #OlympicsOpeningCeremonies. On the other hand, the network comprising of users who invoke #PhDchat appears to represent something more than a spontaneous gathering or information-exchange event, as these individuals gather, self-organize, host a synchronous discussion, and repeatedly return to the network over the one-month period examined.

It is important to note that in our description of community thus far, there is no requirement that the people who gather under its auspices make an ongoing commitment to doing so. Nonetheless, the information available in our dataset does not reveal participation patterns over a significant amount of time. Communities in general experience some amount of turnover, but the dynamic and spontaneous nature of Twitter makes it possible that #PhDchat could share no common members day after day. Twitter is simultaneously a synchronous and an asynchronous communication medium: members may read messages from others immediately and respond as if they were in the same room or review the digest days later. Not only do the members who are communicating at any one time fluctuate rapidly, but the users who may be interacting more infrequently may respond or simply passively observe long after an individual has posted his or her first and only #PhDchat tweet. In this constantly shifting environment, it appears imprecise to say that the network is made up of individuals, because the construct in which they are gathering is made up of instances of connection assembled through a simple classification, the hashtag. Since this social structure is made of ad hoc connections rather than established norms and procedures, "membership" may be granted to any individual who chooses to tap into the #PhDchat stream by supplying or consuming information. This means that the size, shape, and composition of the group are in continuous states of emergence and change. Within this context, Twitter-based networks and communities may have short and evolving memories.

Implications

The analysis presented above created a snapshot of the #PhDchat network during the time of the study. Via an analysis of the interactions between members of the #PhDchat network, we drew conclusions about the nature of this group and its characteristics. At any one time, #PhDchat represents the desires and needs of its members, and its ability to disseminate information is key to its mechanisms for sustaining itself. The observed attributes of #PhDchat, such as the quality, scope, and level of discourse and interaction, can be extrapolated into implications for other online learning and support groups. The phenomenon of the emergent social network community provides insight into the ways in which learners may organize in order to facilitate their own learning.

#PhDchat demonstrates that disparate users can come together with little central authority in order to create their own communal space. The organization is democratic in that participation is relatively open, requiring one only to use the hashtag in order to participate. However, like any other social structure, the network may fragment into

clusters of individuals who interact with one another for different purposes. At the centre, the largest and most connected group is a gathering of frequently contributing users who share links, answer questions, and participate in regular discussions.

One ingredient of an emergent online learning network is its users' willingness to continue creating it. Because #PhDchat, as a stream of tagged tweets, would fade from existence altogether if users stopped including the hashtag in their message, each time an individual types "#PhDchat" at the end of a tweet they are confirming the validity of the network. This willingness might come from the perceived utility of the network. If the network was not providing something of value to its members, it would not exist. Therefore, the strength and limitation of #PhDchat is its transience. It is created and defined by its parts and can change to fit the needs of its members at any time.

Conclusion

In this paper, we sought to answer the question: What are the structure and characteristics of the network that has formed around the #PhDchat hashtag on Twitter? We found that #PhDchat is made of thousands of users who contribute at differing levels. A small core of users participates frequently and attends the weekly discussions, and many users are connected to the community through interactions with this group. There is evidence in the data set to suggest that participants share resources with, offer advice, and provide social and emotional support to one another. Much of the communication is directly related to the process of obtaining a PhD. The use of hashtags is popular within the group of core users and many include hashtags in tweets that link them to other online communities suggesting that they may participate in multiple support networks. The community's status is facilitated by the presence of few barriers to entry and by Twitter's fast pace. As a result, the community is in a continuous state of emergence and change.

This study faces a number of limitations. First, the data we collected allow us to observe user behavior, but not intent. To examine intent, motivations, and reasoning behind the data, we need to use different methodologies and data collection techniques. Second, while the network visualizations and statistics lend some insights into the structure of the community, we do not claim that they provide a complete picture.

Congregation around education-related hashtags such as #edchat, #edtech, #BCed, and #cdnpse or course-related hashtags provides unique research opportunities. For example, #PhDchat is a consistently active hive of contributions and #PhDchat participants generate a large amount of information outside of Twitter (e.g., blog posts, community wiki posts). These spaces may hold evidence pertaining to the knowledge-building that is taking place in this network or even reveal other clusters of members. Future studies may expand the investigation of #PhDchat into these artifacts. Furthermore, alternative data collection methods (e.g., interviews and focus groups) may yield additional information about what the community is achieving and how it affects participant experiences. This kind of research may also endeavor to understand the motivation to participate in and the rewards that one derives from such a

community. Such insights will generate a richer picture of the network and the ties that exist between network members.

Of particular interest to researchers studying emerging online environments may be the lifecycle and changing dynamics of the network over time. Since #PhDchat includes doctoral students, its members may naturally evolve from new students to experienced students to working academics. The changing membership and shifting professional roles may affect the dynamics of the group and may have implications for the functions served by this community. The ease with which users can leave the network may imply that it is fragile, but the fact that users can join it with the same ease may suggest that low barriers to entry may sustain the network. A longitudinal study of tweets, users, and wider network activity could hold clues to what draws members into and repels them away from communities like #PhDchat over time.

Though this community may be of particular interest to educational researchers, the groups are intangible, making them difficult to study. SNS are third-party; for-profit ventures and collecting information responsibly can be a challenge. Furthermore, the social nature of the medium adds a complex layer of interpersonal dynamics to the context of the study. More research is needed to create a model for understanding emergent social network communities and make recommendations for how such learning networks can be more effectively studied, analyzed, and understood by researchers.

References

Aspden, E., & Thorpe, L. (2009). "Where Do You Learn?": Tweeting to Inform Learning Space Development. EDUCAUSE Review. Retrieved February 21, 2013, from http://www.educause.edu/ero/article/where-do-you-learn-tweeting-inform-learning-space-development

Blankenship, M. (2011). How Social Media Can and Should Impact Higher Education. Education Digest, 76(7), 39-42.

Boyd, Danah M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.

Chen, M. (2011). Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. Computers in Human Behavior, 27(2), 755-762.

Coiffait, F., Bartlett, K., Houghton, D. J., & Condie, J. (in press). Social media and clinical psychology: Friend or foe? Clinical Psychology Forum.

Couros, A. (2010). Developing personal learning networks for open and social learning. In Veletsianos G. (Ed.), Emerging Technologies in Distance Education (p. 109-128). Edmonton, AB: Athabasca University Press.

DeAndrea, D. C., Ellison, N. B., LaRose, R., Steinfield, C., & Flore, A. (2012). Serious social media: On the use of social media for improving students' adjustment to college. The Internet and Higher Education, 15(1), 15-23.

Doll Jr., W. E. (2009). Complexity Theory. In Encyclopedia of the Social and Cultural Foundations of Education. Sage Publications.

Dron, J., & Anderson, T. (2009). How the Crowd Can Teach. In Handbook of Research on Social Software and Developing Community Ontologies (pp. 1-17). Hershey, PA: Information Science Reference.

Duggan, M., & Brenner, J. (2013). The Demographics of Social Media Users - 2012 (pp. 1-14). Washington, D.C. Retrieved from http://pewinternet.org/Reports/2013/Social-media-users/The-State-of-Social-Media-Users.aspx

Ebner, M., Lienhardt, C., Rohs, M., & Meyer, I. (2010). Microblogs in Higher Education - A chance to facilitate informal and process-oriented learning? Computers & Education, 55(1), 92-100.

Ferriter, W. (2010). Why Teachers Should Try Twitter. Educational Leadership, 73-74.

Forte, A., Humphreys, M., & Park, T. (2012). Grassroots Professional Development: How Teachers Use Twitter. In Sixth International AAAI Conference on Weblogs and Social Media (pp. 106-113).

Gerstein, J. (2011). The Use of Twitter for Professional Growth and Development. International Journal on E-Learning, 10(3), 273-276.

Gil de Zúñiga, H., Nakwon, J., & Valenzuela, S. (2012). Social Media Use for News and Individuals' Social Capital, Civic Engagement and Political Participation. Journal of Computer-Mediated Communication, 17(3), 319-336.

Glynn, C. J., Huge, M. E., & Hoffman, L. H. (2012). All the news that's fit to post: A profile of news use on social networking sites. Computers in Human Behavior, 28(1), 113-119.

Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Learning, Teaching, and Scholarship in a Digital Age. Educational Researcher, 38(4), 246 -259.

Gruzd, A., & Haythornthwaite, C. (2011). Networking online: cybercommunities. In The SAGE Handbook of Social Network Analysis (pp. 167-179). London: Sage Publications.

Herring, S. C. (2008). Computer-Mediated Discourse. In D. Tannen, D. Schiffrin, & H. Hamilton (Eds.), Handbook of Discourse Analysis. Oxford: Blackwell.

Holmes, K., Preston, G., Shaw, K., & Buchanan, R. (2013). 'Follow' Me: Networked Professional Learning for Teachers. Australian Journal of Teacher Education, 38(12).

Honey, C., & Herring, S. C. (2009). Beyond Microblogging: Conversation and

Collaboration via Twitter. In HICSS '09. 42nd Hawaii International Conference on System Sciences (pp. 1-10).

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why We Twitter: Understanding Microblogging Usage and Communities. In Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07 (pp. 56-65). New York, NY: ACM Press.

Jenkins, H., Clinton, K., Purushotma, R., Robison, A. J., & Weigal, M. (2006). Confronting the Challenges of Participatory Culture: Media Education for the 21st Century (pp. 1-72). Retrieved from http://digitallearning.macfound.org /site/c.enJLKQNIFiG/b.2108773/apps/nl/content2.asp?content_id= {CD911571-0240-4714-A93B-1D0C07C7B6C1}&notoc=1

Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. Journal of Computer Assisted Learning, 27(2), 119-132.

Kassens-Noor, E. (2012). Twitter as a teaching practice to enhance active and informal learning in higher education: The case of sustainable tweets. Active Learning in Higher Education, 13(1), 9-21.

Kaya, T. (2010). CUNY Social Network Mixes Scholarship With Facebook-Style Friendship. The Chronicle of Higher Education.

Lalonde, C. (2011). The Twitter experience: The role of Twitter in the formation and maintenance of personal learning networks. MA Thesis. Royal Roads University.

Lin, M.-F. G., Hoffman, E. S., & Borengasser, C. (2013). Is Social Media Too Social for Class? A Case Study of Twitter Use. TechTrends, 57(2), 39-45.

McInnerney, J., & Roberts, T. (2004). Online learning: Social interaction and the creation of a sense of community. Journal of Educational Technology, 7(3), 73-81.

Parker, A. (2011, June 10). Twitter's Secret Handshake. The New York Times. Retrieved from http://www.nytimes.com/2011/06/12/fashion/hashtags-a-new-way-for-tweets-cultural-studies.html

Petrilli, M. J. (2011). All A-Twitter about education: improving our schools in 140 characters or less. Education Next, 11(4), 90-91.

Reed, P. (2013). Hashtags and retweets: using Twitter to aid Community, Communication and Casual (informal) learning. Research in Learning Technology, 21. Retrieved from http://www.researchinlearningtechnology.net/index.php/rlt/article /view/19692/html

Rheingold, H. (2010). Attention, and Other 21st-Century Social Media Literacies. EDUCAUSE Review, 45(5), 14-24.

Riel, M., & Polin, L. (2004). Online learning communities: Common ground and critical differences in designing technical environments. In Designing for Virtual Communities in the Service of Learning (pp. 16-50).

Thackray, L. (n.d.). PhD Chat Wiki. Retrieved April 06, 2013, from http://phdchat.pbworks.com/w/page/33280234/PhD%20Chat

Thompson, C. (2008). Brave New World of Digital Intimacy. The New York Times.

Veletsianos, G. (2012). Higher Education Scholars' Participation and Practices on Twitter. Journal of Computer Assisted Learning, 28(4), 336-349.

Veletsianos, G. (2013). Open Practices and Identity: Evidence from Researchers and Educators' Social Media Participation. British Journal of Educational Technology, 44(3), 639-651.

Veletsianos, G. & Kimmons, R. (2012). Assumptions and Challenges of Open Scholarship. The International Review Of Research In Open And Distance Learning,13(4), 166-189.

Veletsianos, G., Kimmons, R., & French, K. (2013). Instructor experiences with a social networking site in a higher education setting: Expectations, Frustrations, Appropriation, and Compartmentalization. Educational Technology, Research and Development, 61(2), 255-278.

Welcome to the #loveHE campaign page! (2010). Times Higher Education. Retrieved April 15, 2013, from http://www.timeshighereducation.co.uk/news/welcome-to-the-lovehe-campaign-page/410669.article

Wodzicki, K., Schwämmiein, E., & Moskaliuk, J. (2012). "Actually, I Wanted to Learn": Study-related knowledge exchange on social networking sites. The Internet and Higher Education, 15(1), 9-14.

Young, J. R. (2010). Teaching with Twitter: Not for the faint of Heart. Education Digest, 75(7), 9-12.

Appendix 1

Detailed information on the algorithms used to plot graphs

---

[1] The only exception is #lovehe (or "love higher education"). This hashtag represents a cause/campaign, started by Times Higher Education, a UK-based publication, in March of 2010 to highlight the positive aspects of higher education (Times Higher Education, 2010).

[2] The reader should note that "thesis" used in the UK context is the equivalent of "dissertation" in the North American context.

---

[i] Figure 5: A graph of users and mentions plotted using the Harel-Koren Fast Multiscale layout algorithm. The users were grouped by cluster using the Clauset-Newman-Moore cluster algorithm. Then, the graph was laid out with groups in grid, with major connections combined visually.

[ii] Figure 6: A graph of users and mentions plotted using the Harel-Koren Fast Multiscale layout algorithm. The users were grouped by cluster using the Clauset-Newman-Moore cluster algorithm. Connections between groups were not combined or laid out in a grid.

[iii] Figure 7: A graph of users and mentions with only users that appeared more than once in the dataset. As before the users were broken into groups using grouped by cluster using the Clauset-Newman-Moore cluster algorithm and laid out with the Harel-Koren Fast Multiscale layout algorithm.