

#ncpr2012 Social Media Campaign Report

Peter M Crosta*

Community College Research Center
Teachers College, Columbia University
`peter.crosta@columbia.edu`

July 2, 2012

1 Overview

The National Center for Postsecondary Research (NCPR) held its final conference on remediation at Teachers College, Columbia University on June 21-22, 2012. The Community College Research Center's (CCRC) communications team used the social media tool **Twitter** to update followers on important findings from the conference. Other conference members joined the online discussion by using the hash tag #ncpr2012. This report summarizes the extent of tweeting and gauges the general sentiment of the conference as told by the Twittersverse.

2 Who, when, and how much

How many tweets?

A search on twitter for #ncpr2012 from June 21, 2012 to June 22, 2012 identified 405 tweets. There were 82 retweets and thus 323 unique messages sent. More activity was recorded on the 21st (227 tweets) then on the 22nd (178 tweets).

*Ideas provided by NCPR debrief group

Who tweeted?

Table 1: Who tweets

	Freq		Freq
CCRC	157	Clackamas College	1
Andrew Moore	51	Constance	1
MDRC	47	Christopher Cabaldon	1
Dr. Karen Stout	22	Brooke Barrett	1
Alsace Rodgers	19	Eloy Ortiz Oakley	1
The Hechinger Report	13	Kat Winkey	1
Ann Person	9	Katrina Reichert	1
Bruce Vandal	6	Lauren Kaplan	1
Dean Dad	6	MDC	1
Kresge Education	5	MONEY on College	1
Sara Goldrick-Rab	5	Paula Krebs	1
Achieving the Dream	4	Reading STEMulates	1
Michael Hirsch	4	Sarah Walker	1
bleuguy	4	Parents Union - TX	1
Michael Collins	4	Tom Hilliard	1
DEI	2	Tom Rock	1
Ed Bowling	2	WGUTexas	1
Mandy Zatynski	2	Adam Hall	1
Sharon Morrissey	2	Cathy Grimes	1
Christopher Conzen	2	Clare Bertrand	1
Derek Price	2	Eduard Matamoros	1
John H Rooney III	2	Judith Burstyn	1
Laura Hope	2	kim clark	1
Ron P. Muriera	2	Meghan Oster	1
navjeet singh	2	mobilediscoapp	1
CCBC	1	Mamadou Ndiaye	1
Ellen S. Boyne	1	Andrea OC	1

The majority of tweets came from CCRC. However, several other conference attendees were active at the time. Table 1 lists the 54 unique tweeters and the number of times they tweeted during the two-day event. Note that most users only contribute one comment as very few users dominate the con-

versation. Future campaigns could make more use of @replies, @messages, and direct questions to increase engagement.

When did they tweet?

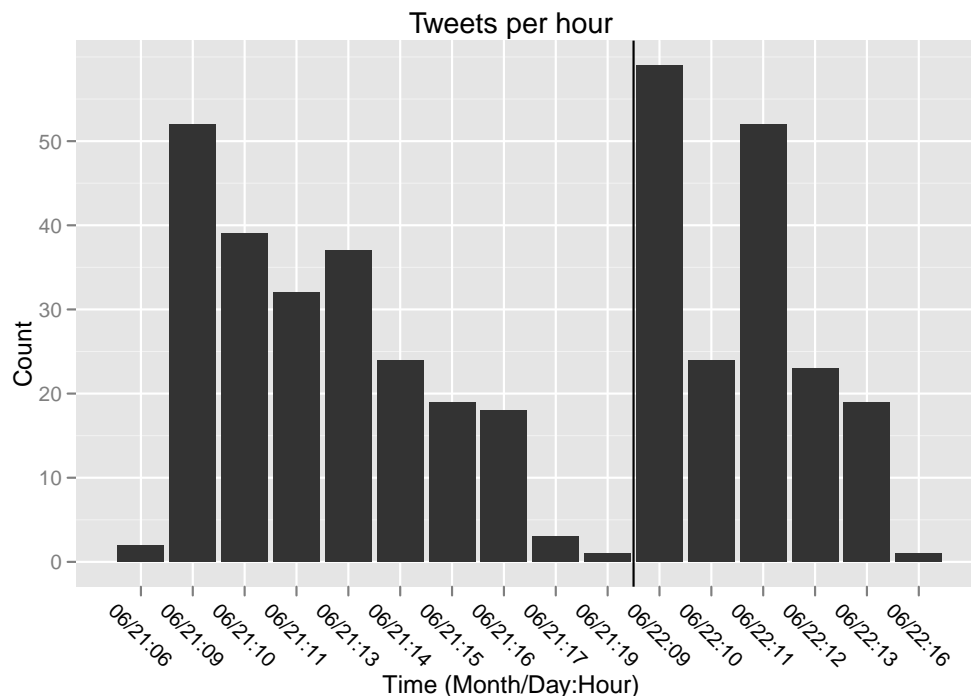


Figure 1: Tweet frequency

As shown in Figure 1, Twitter activity was greater during the morning sessions than the afternoon sessions, with a slow decline over the course of the first day. The 9am and 11am sessions on the second day saw a spike in activity, but most hours saw only about 20 tweets.

Figure 2 illustrates who was tweeting and when they were doing it during the conference. Perhaps predictably, tweet activity by MDRC spiked during the 9am hour on the 22nd, when MDRC researchers were reporting initial findings from research on Accelerated Study in Associate Programs (ASAP). We might categorize users into two different types. One group periodically tweets, leaving lots of time between individual messages. The second type

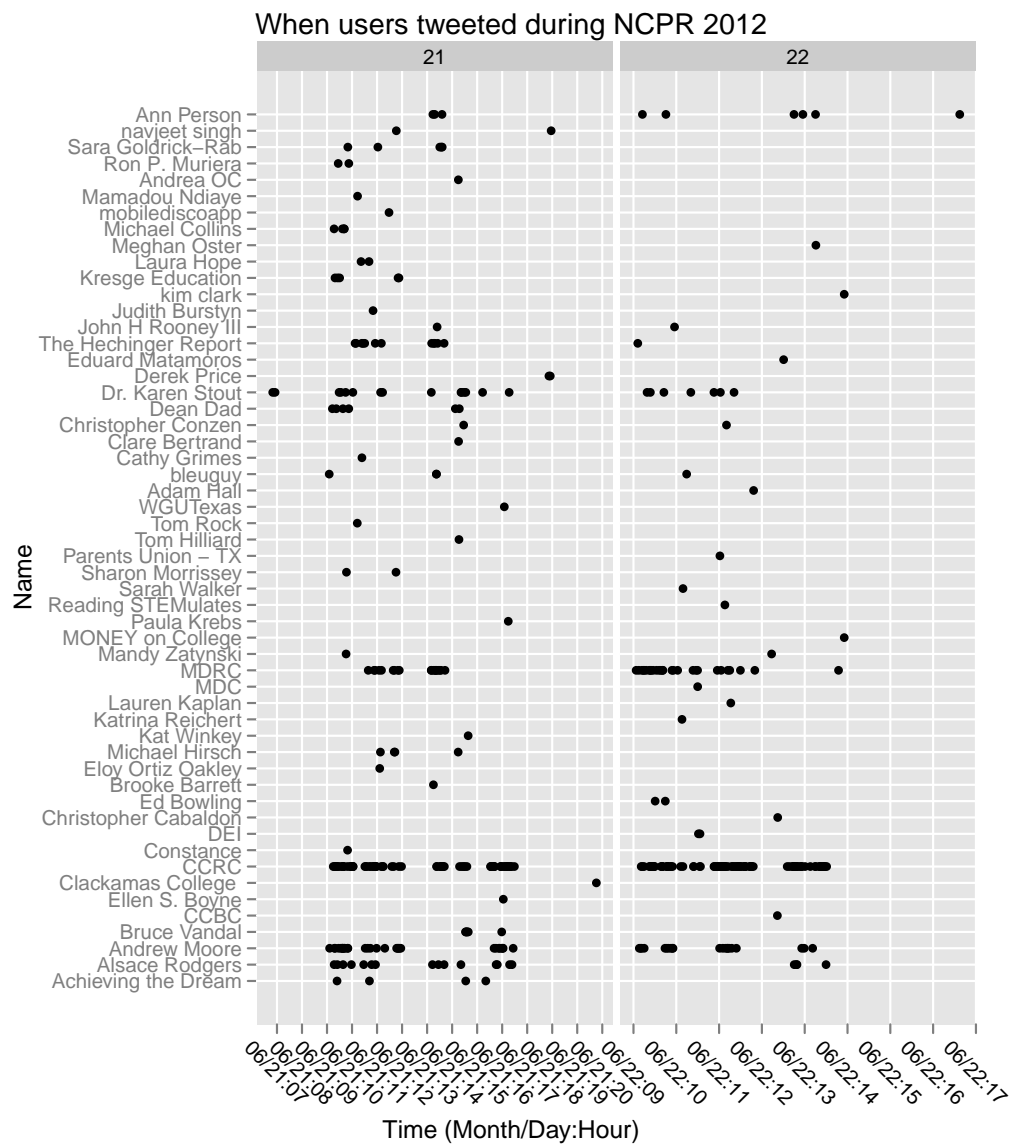


Figure 2: Who tweeted when

of user generated a clumping pattern. These users have higher tweet counts and would fire several tweets in rapid succession.

3 Sentiment analysis

To gauge the overall tone of conference-related tweets, we conduct a short sentiment analysis using two strategies. The first and most basic strategy is a sentiment analysis that involves comparing the text in tweets to banks of positive and negative words. The general sentiment score is the difference in the number of positive words and the number of negative words¹. The second and somewhat more sophisticated analysis employs an algorithm that has been trained to recognize positive and negative text.

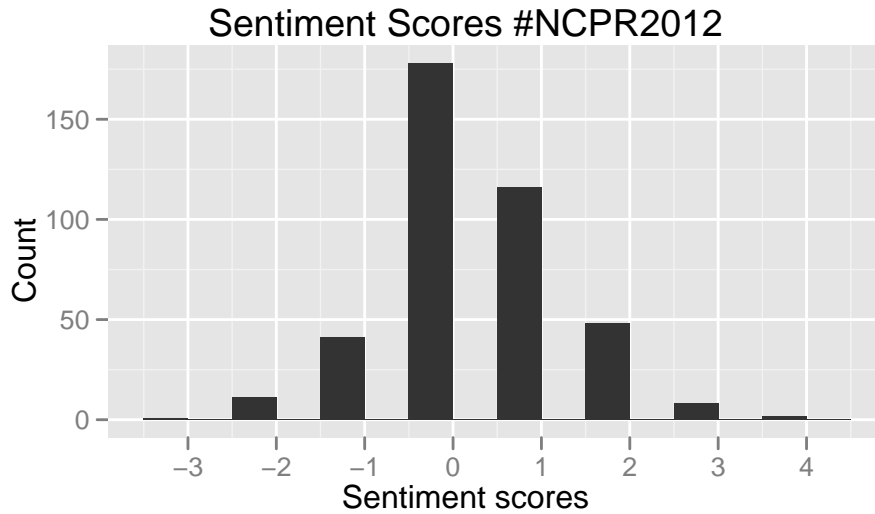


Figure 3: Histogram of sentiment scores.

As an example of tweets with positive and negative text, consider the following highest-rated tweet written by CCRC:

Thanks all for a fantastic conference. Really inspiring group of people who are dedicated to seeing their students succeed!!
#ncpr2012

The lowest-rated tweet was written by **The Hechinger Report**:

¹See Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA

Remedial ed: should we kill it, smash it or blow it up, McClenney asks?#ncpr2012

Figure 3 shows the distribution of sentiment scores. Most tweets receive a score of zero. They either have words that cancel each other out or have words that do not appear on either list. The mean sentiment by the first method (higher is better) is 0.44; Excluding tweets from CCRC the sentiment is 0.44. This means that the tweets were, on average, more positive than negative.

Using the second sentiment method (Naive Bayes), we are able to label each tweet as Negative, Neutral, or Positive. The classification shown in Table 2. Again we see that tweets are more positive than negative.

Table 2: Tweet sentiment breakdown

	Freq
Negative	59
Neutral	37
Positive	309

It matters who tweets

It is important to remember that all tweets are not created equal. Some Twitter users have tens of thousand of followers, reaching a very wide audience, while others have only a handful of followers. To take into account the importance of a tweet, we weight the sentiment score by the number of followers that each user has. Weighting by user importance reduces the average sentiment score to 0.374, or 0.292 when omitting CCRC tweets. These findings suggest that overall positive tweet sentiment is being driven by CCRC tweets and tweeters of higher importance are generally less positive than tweeters of lower importance.

4 What did they say?

A full list of tweet text can be obtained by e-mailing the author. One way to get a big picture of the conference is to create a word cloud based on the text of the tweets. This can be seen in Figure 4.

