#Parenting Projects: Using Twitter to Understand Mothering and Fathering

Joyce Lee, Garrett Pace, Alex Lu, Cheryl Hershey, and Tawfiq Ammari

Methods Hour
April 5, 2019
Agenda

● Introduction and background
● Use of social media to study parenting
● Our project using Twitter data to study parenting self-disclosure online
● Challenges and opportunities using Twitter data to engage in parenting research & related discussions
Introduction: Parents on Social Media

- Today’s parents, especially those with young children, turn to the web and social media for parenting information.
- Of them, 74% who use social media receive support.
- Mothers more likely than fathers to seek parenting information and get social support from friends.
- Parents use various social media platforms.
  - More recently, parents turning to YouTube for children’s content (Smith, Toor, & Kessel, 2018)
  - (Duggan, Lenhart, Lampe, & Ellison, 2015)
Parents Use a Range of Social Media Platforms; Facebook Tops the List

Among all internet users, the % of parents who use each social media platform

<table>
<thead>
<tr>
<th>Platform</th>
<th>Parents</th>
<th>Mothers</th>
<th>Fathers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>74%</td>
<td>81%</td>
<td>66%</td>
</tr>
<tr>
<td>Pinterest</td>
<td>28</td>
<td>40</td>
<td>15</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>27</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Instagram</td>
<td>25</td>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>Twitter</td>
<td>19</td>
<td>23</td>
<td>27</td>
</tr>
</tbody>
</table>

* Differences between mothers and fathers are statistically significant

Social Media is One of Many Sources for Parenting Advice and Information

Among all parent social media users, the % who have done the following on social media over the previous month...

- Found parenting info while looking at social media content:
  - All parents: 59
  - Mothers: 66
  - Fathers: 48

- Received social/emotional support on a parenting issue:
  - All parents: 42
  - Mothers: 50
  - Fathers: 28

- Asked parenting questions:
  - All parents: 31
  - Mothers: 36
  - Fathers: 24

* Differences between mothers and fathers are statistically significant

Source: Pew Research Center surveys, Sept. 11-14 and 18-21, 2014. N=1,597 internet users ages 18+. The margin of error for all internet users is +/- 2.9 percentage points. Parents in this survey were defined as those with children under age 18.

PEW RESEARCH CENTER
Use of Social Media to Study Parenting

- Prior studies have used Facebook, Reddit, and parenting blogs to study parenting.

- Our rationale for choosing Twitter to study mothering and fathering:
  - Open and public source of data
  - Parenting beliefs can be observed naturalistically, potentially addressing social desirability bias
  - Parents turning to Twitter for parenting information and social support

(Ammari et al., 2018; Bartholomew et al., 2012; Schoenebeck, 2013)
Our Project Using Twitter to Study Mothering and Fathering

- Sample focused on stay-at-home mothers and stay-at-home fathers
- Both stay-at-home mothers and fathers have increased in number in recent decades
  - Fathers: 1.1 million in 1989 to 2.0 million in 2012
- A group that possibly defies traditional gender norms
- Research has shown that stay-at-home parents experience social isolation and turn to social media for emotional support

(Cohn & Caumont, 2014; Lee & Lee, 2018; Livingston, 2014, 2018)
Our Project Using Twitter to Study Mothering and Fathering (cont.)

- Our three studies are largely exploratory in nature.
- General theme focused on comparing stay-at-home fathers’ and stay-at-home mothers’ Tweets for similarities and differences.
- Three main studies for this presentation:
  1. Modeling parenting topics
  2. Geocoding user-defined location
  3. Qualitative coding of spanking Tweets
Data Collection: Stay-At-Home Parents’ Tweets

- Queried hashtags (e.g., #stayathomedad, #stayathomemom) for 30 days using Twitter API
- Used regular expressions on Twitter profiles to select self-identifying stay-at-home parents
- Utilized this seed dataset to engage in snowball sampling of followers for several iterations
- Cleaned up user data (e.g., eliminate duplicates)
- Total 697 unique father and 4,103 mother users
Husband to a most awesome woman. Father of 3 amazing boys and 1 girl. Foster Dad. Stay at Home Dad. Beekeeper, I’ve been resisting since I was born.

Runner of 5k to 100 miles, cross country & track coach, stay-at-home father of 3 boys & 2 girls, jack-of-all-trades. I ferment things. Live with thadditude.

Just the every day occurences in the life of a stay at home mom with her toddler and newborn ;)

Stay at home mom to three girls, pixel artist(GameMaker)believes that good people exist despite overwhelming evidence to the contrary. MomSquad
Apply regular expressions

Seed mothers (n = 37)
Seed fathers (n = 15)

First iteration of followers (n = 209 fathers)
First iteration of followers (n = 761 mothers)

Data cleaning (e.g. exclude duplicates, protected or dead accounts)

Second iteration of followers (n = 9,462 fathers)
Second iteration of followers (n = 35,130 mothers)

Query hashtags (n = 3,087 unique users)

Total fathers (n = 697)
Total mothers (n = 4,103)
Data Collection: Stay-At-Home Parents

● Obtained stay-at-home parents’ tweets using Twitter Resting API
  ○ Can get up to 3,200 most recent tweets
● Total 936,985 tweets from stay-at-home fathers
● Total 4,694,124 tweets from stay-at-home mothers
● Data cleaning involved eliminating punctuations, encoding emojis and symbols, lowercasing; collaboration with data scientists helpful
A single tweet provides a ton of information or metadata

- Screen name, user-defined location, language, number of followers, coordinates in which tweet is being posted
- Twitter is constantly updating metadata features
Project #1: Modeling Parenting Topics

● Purpose was to model parenting topics stay-at-home fathers and stay-at-home mothers discuss on Twitter
  ○ Comparing similarities and differences
● Each user’s tweets was a corpus, a collection of texts
● Used Latent Dirichlet Allocation (LDA) for topic modeling
● Used Word2Vec word embeddings to help differentiate between topics stay-at-home fathers and stay-at-home mothers discuss on Twitter
Project #1: Modeling Parenting Topics (cont.)
Results from Project #1

- Examples of shared topics amongst stay-at-home fathers and stay-at-home mothers:
  - Parenting experiences
  - Parenting philosophy
  - Do-It-Yourself projects and crafts
  - Giveaways, shopping, and monetization
Results from Project #1 (cont.)

- Topics unique to fathers and mothers:

<table>
<thead>
<tr>
<th>Topic #</th>
<th>Topic Name</th>
<th>KTG</th>
<th>Prob distribution</th>
<th>SAHD/SAHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fatherhood advocacy/support</td>
<td>[ dad, kid, thank, par, follow, gre, dadsummit, lov, get, tim, sahd, ad, mom, day, fath, us, w, lik, amazonfamily, fam]</td>
<td>0.02</td>
<td>SAHD</td>
</tr>
<tr>
<td>2</td>
<td>Brewing</td>
<td>[ badg, earn, gard, beer, level, al, brew, ip, drink, day, new, superst, plant, photo, post, lik, today, see, gre, ti]</td>
<td>0.01</td>
<td>SAHD</td>
</tr>
<tr>
<td>3</td>
<td>Weight loss/exercise</td>
<td>[cal, fitbit, step, travel, mil, minut, lbs, exerc, diary, fitstats, myfitnesspalburn, food, goal, fitst, weigh, lost, sint, far, walk, mp]</td>
<td>0.03</td>
<td>SAHM</td>
</tr>
</tbody>
</table>

Table 1. This table shows the three unique topics that were discovered from the independent LDA models trained on SAHD tweets vs. SAHM tweets.
Project #2: Geocoding Stay-At-Home Parents’ Self-Defined Locations

- Purpose was to employ user-defined location to plot where stay-at-home parents are and prevalence of tweets from those locations
- User-defined location available on Twitter profiles

Joyce Y. Lee
@joyceyaeunlee
PhD student at University of Michigan. Interested in the role of the father in child development and wellbeing, technology-based interventions, & data science.

Ann Arbor, MI
joyceyee.com
Joined February 2009
Project #2: Geocoding Stay-At-Home Parents’ Self-Defined Locations (cont.)

- Methods primarily involved combining Twitter user-defined location metadata and Google API to identify latitudes and longitudes
- Deleted parents without user-defined location
- Plotting data points via data visualizations tools in R (i.e., ggplot2 and ggmap)
Stay-At-Home Fathers \((n = 553)\)
Narrowing in on U.S. Stay-At-Home Fathers
Proportion of Tweets by U.S. Stay-At-Home Fathers
Stay-At-Home Mothers \((n = 2,165)\)
Narrowing in on U.S. Stay-At-Home Mothers
Proportion of Tweets by U.S. Stay-At-Home Mothers
Project #3: Content Coding Stay-At-Home Parents’ Spanking Tweets

- Purpose was to examine stay-at-home mothers’ and stay-at-home fathers’ tweets concerning discipline, with a focus on spanking
- Involved combining both data science and qualitative coding methods
  - Used regular expressions to identify tweets that mention “spanking” and its variations (e.g., spanked)
  - Qualitative coding the tweets that were identified to look for emerging themes
Project #3: Content Coding Stay-At-Home Parents’ Spanking Tweets (cont.)

- Also, looked for synonyms to spanking (e.g., beat, punch, slap, tap, whoop)
- Overall, found a very small set of tweets which suggests that stay-at-home parents may not readily discuss this parenting topic on Twitter
  - 69 tweets from fathers
  - 240 tweets from mothers
Results from Project #3

● Anti-spanking tweets common:
  ○ “If you spank my child, I will reign down my wrath on you. Other than that, I’m a pretty #chillaxdad” by @realsahd

● Substantial proportion of pro-spanking tweets:
  ○ “I believe there’s a difference between spanking & abusing and I see nothing wrong w/ the former as a last resort” by @savvymommy

● Information tweets quite common as well:
  ○ “RT children who are spanked have lower IQ’s compared to children who are not spanked according to new research. [link to article]” by @sahmlife
Challenges and Opportunities of Using Twitter Data in Parenting Research

● Challenges
  ○ Ethics: Consent? Compensation?
  ○ Time when data collected and study published

● Opportunities
  ○ Lots of free data to collect (text, network, metadata)
  ○ Space for creativity and innovation
  ○ Twitter likes academics
Challenges and Opportunities of Using Twitter Data: Discussion Questions

● What other challenges and opportunities do you see when considering using Twitter data to conduct social science research?
● How might you use Twitter data in your own research and what research questions might you consider answering?
● What feedback might you anticipate from journal reviewers and how might we best address them?
References

- Ammari et al., 2018: https://doi.org/10.1145/3173574.3174063
- Cohn & Caumont, 2014: https://www.pewresearch.org/fact-tank/2014/04/08/7-key-findings-about-stay-at-home-moms/
- Lee & Lee, 2018: http://dx.doi.org/10.1037/men0000079
- Schoenebeck, 2013: https://pdfs.semanticscholar.org/9ea1/c4f83fdf50ed904c6797dcea91f0452c75e6.pdf