Understanding gender and skills gaps beyond Wikipedia (Part 1)

Aaron Shaw, Northwestern University, aaronshaw@northwestern.edu
Krystle Chung, wikiHow, krystle@wikihow.com
Eszter Hargittai, Northwestern University, eszter@northwestern.edu
Motivation

- Comparison: Few surveys beyond WP
- Understand wikiHow (build on 2012 survey)
- Address data quality & sampling concerns
- Extend gender & skills gaps research
wikiHow at a Glance

13,000 edits/week • 600 first edits/week • 1400 total editors/week
2012 wikiHow Community survey

- 122 respondents
- 44% male, 56% female
- 76% 25 or younger
- Females outnumber males among ≤25
- The older the contributor, the more likely to be male

See http://youtube.com/watch?v=zXsUNcITjQ4
The Wikipedia Gender Gap Revisited: Characterizing Survey Response Bias with Propensity Score Estimation

Benjamin Mako Hill, Aaron Shaw

Published: June 26, 2013 • DOI: 10.1371/journal.pone.0065782

Abstract

Opt-in surveys are the most widespread method used to study participation in online communities, but produce biased results in the absence of adjustments for non-response. A 2008 survey conducted by the Wikipedia Foundation and United Nations University at the 37th IETF (Internet Engineering Task Force) conference found that 76% of survey respondents were male. Using a random sample of contributors to the English Wikipedia, we estimate that the survey response bias could lead to a 0.8% overestimation of the male/female ratio for Wikipedia contributors. The results are consistent with a recent study from a large internet survey. These findings highlight the importance of adjusting for non-response when conducting surveys on the internet.
Gender gap vs. Internet skills gap
Methods (our plan)

- Build a random sample of ~600 wikiHow editors
- Recruit via talk page messages, follow-up emails
- Distribute an online survey using Qualtrics
- Match survey results with edit history
- Analyze and release anonymized dataset (later…)}
Survey questions & measures

- Demographic attributes (screen out respondents <18)
  - gender, edu, employment, country, race, income
- Attitudes about wikiHow
- Confidence editing wikis
- Wikipedia editor?
- Internet skills index
- Trust/caution scale
- Attention test

(we use standard, validated, and widely adopted measures wherever possible)
### Internet skills index

How familiar are you with the following computer and Internet-related items? Please choose a number between 1 and 5 where 1 represents “no understanding” and 5 represents “full understanding” of the item.

<table>
<thead>
<tr>
<th>No understanding</th>
<th>Little</th>
<th>Some</th>
<th>Good</th>
<th>Full understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced search</td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
</tr>
<tr>
<td>Cache</td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
</tr>
<tr>
<td>PDF</td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
</tr>
<tr>
<td>Phishing</td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
</tr>
<tr>
<td>RSS</td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
</tr>
</tbody>
</table>

Sampling and Survey Distribution

- Stratify 1400 weekly editors into 3 groups.
- Randomly sample within strata.
- ≤20 questions.
- Repeat weekly, remove anyone asked in previous week.
- Started July 8 2015 (about 60 valid responses so far…)
Preliminary Results: Overview

- 59 valid responses so far (out of 693 invited; 8.5%)
  - 28 are 18 or older (31 under 18)
  - answered attention question correctly
  - vast majority are “high” contributors
- Age:
  - range up to 62 (most between 21 and 40)
- Most have at least some college or university education
- Most working part & full time
- Most have edited Wikipedia at least a few times.
Preliminary Results: Gender

● About half of respondents are female
  ○ 11 female, 17 male
  ○ Consistent with prior, informal survey!

● Evenly split among editors with 5+ edits, 7:1 male ≤4 edits
Preliminary Results: Internet skills

- Most respondents have above average Internet skills.
- Nobody with low (<2.5) Internet skills has responded.
- Similar to prior findings from Hargittai & Shaw (2015)!
Gender gap, Internet skills & Edits

Gender, Internet Skills, and Contributions to wikiHow

- Total edits on wikiHow
- Internet Skills (Possible range from 1–5)

Gender
- Female
- Male
Preliminary takeaways...

(active) wikiHow editors are mostly:

- Young (about half under 18; vast majority under 40)
- US-based, highly educated, (partly) employed.
- Highly skilled Internet users.

No sign of gender/skills gap among editors (yet)
Future work on this project...
Trust/caution and participation

- Do people who contribute more tend to exhibit more generalized trust?
- Compare with previous findings from Wikipedia and population studies.
So, we’ll be back soon with...

- More data!
- More nuanced and complete analysis.
- More insights into how gender and skill gaps may shape participation in wikis.