Digital Literacy and Sampling: MTurk Bad

Kevin Munger, Mario Luca, Jonathan Nagler, Joshua Tucker

Penn State University and Princeton University

June 22, 2019
Online media effects are rapidly changing—how do we keep up?
Overview

- Online media effects are rapidly changing—how do we keep up?

- Online media effects are uniquely heterogeneous—how do we generalize, how do we study the correct populations?
Overview

- Online media effects are rapidly changing—how do we keep up?
- Online media effects are uniquely heterogeneous—how do we generalize, how do we study the correct populations?
- Our intuitions about social media can be actively misleading—how do we adjust?
Overview

- Online media effects are rapidly changing—how do we keep up?

- Online media effects are uniquely heterogeneous—how do we generalize, how do we study the correct populations?

- Our intuitions about social media can be actively misleading—how do we adjust?

- Case study: clickbait!
Null Results from First MTurk Study

- Tried again: made the subject matter more topical, added a placebo condition
Null Results from First MTurk Study

- Tried again: made the subject matter more topical, added a placebo condition

- (After the pilot, we pre-registered the R code we used to analyze all results)
Tried again: shortened the survey, removed “preference for clickbait” questionnaire which could dampen treatment effects
Null Results from Third MTurk Study
Null Results from Third MTurk Study

- Is Mturk the problem?
Null Results from Third MTurk Study

- Is Mturk the problem?

- Many classic (non-digital) experiments replicate on MTurk (Coppock, 2018)
Null Results from Third MTurk Study

- Is Mturk the problem?

- Many classic (non-digital) experiments replicate on MTurk (Coppock, 2018)

- But: MTurk users are all above a certain threshold of digital literacy
Clickbait Effects on the Clickers

NYU Survey
Create Page @Username

Home
About
Events

See more
Promote
Manage Promotions

NYU Survey
Just now

Win up to 500$ with our 5 minutes survey. Give us your opinion and get a chance to win an Amazon voucher of 500$! (more info on the website)

WIN UP TO 500$ WITH 5 MIN SURVEY

Online Survey Software | Qualtrics Survey Solutions
Qualtrics sophisticated online survey software solutions make creating online surveys easy. Learn more about Research Suite and get a free account today.

NYU.QUALTRICS.COM
Null Results from the FB Study

- We got the right sample and didn’t find results
Attrition from Online Samples

False start
Roll Off
New Tab
DVs

Percentage of Respondents

group
Facebook
MTurk
Null Results from the FB Study

- We got the right sample and didn’t find results. Attrition was non-random and covaried with demographics of interest.
Ages of FB Sample at Attrition Points

Density of Respondents

Age

Finished
New Tab
Other

Age

20 40 60 80
Examine Predictors of Stopping at New Tab

- Combine the data, run a fully interacted model to look at differential effects in the two samples
Effect of Age on Stopping at New Tab: MTurk v Facebook
“Attention Checks” With Digitally Naive Populations

If you could only choose to read one of these stories, which would it be?

- Here's what happened with the Dallas Cowboys this weekend
- Drake and Rihanna are getting back together after a vacation in Area 51
- CNN tweets new response to controversy

Survey taker: always select this option, ignore the other three headlines

Passed attention check MTurk: 82%
Passed attention check FB: 57%
Attention Checks

- Combine the data, run a fully interacted model to predict missing the attention check
Attention Checks

- Combine the data, run a fully interacted model to predict missing the attention check

- Differential effects (similar to above)
Attention Checks

- Combine the data, run a fully interacted model to predict missing the attention check

- Differential effects (similar to above)

  Effect of Age on Missing Attention Check: MTurk v Facebook
Attention Checks

- Combine the data, run a fully interacted model to predict missing the attention check

- Differential effects (similar to above)

  Effect of Age on Missing Attention Check: MTurk v Facebook
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”
- All but 1 MTurker entered a number
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”
- All but 1 MTurker entered a number
- 39 Facebookers entered [eg]:
  - Seventy one years
  - 78 and not senile.
  - 67
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”

- All but 1 MTurker entered a number

- 39 Facebookers entered [eg]:
  - Seventy one years
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”
- All but 1 MTurker entered a number
- 39 Facebookers entered [eg]:
  - Seventy one years
  - 78 and not senile.
Open-response question: “How old are you”

All but 1 MTurker entered a number

39 Facebookers entered [eg]:
- Seventy one years
- 78 and not senile.
Non-Numeric Ages Entered Into Text Box: Digital Dexterity

- Open-response question: “How old are you”

- All but 1 MTurker entered a number

- 39 Facebookers entered [eg]:
  - Seventy one years
  - 78 and not senile.
  - ,67
Older People Using Mechanical Turk

- Qualitative study of older, non-Mturk users: they can’t do basic tasks on MTurk (Brewer, Morris, and Piper, 2016)
Older People Using Mechanical Turk

- Qualitative study of older, non-Mturk users: they can’t do basic tasks on MTurk (Brewer, Morris, and Piper, 2016)
- Modal respondent reported having used the internet for 10 years or more
many participants were not familiar or comfortable with opening content in new tabs/windows....‘How do I get back to the instructions? (P7)’....P3 explained: ‘There’s too many things to remember all at once...One of my complaints about some things on a computer is that, you know, if there’s a bunch of instructions or stuff to know — and you have to open up a box and then if you go back to what you’re working on the box is gone, and you can’t just look up [sic] and reference it. (Brewer, Morris and Piper, 2016)
These barriers, which may seem trivial from a requester’s perspective, significantly affected older adults’ abilities and time required to complete the tasks. And, because of these barriers, some tasks expired or ran out of time before the older adult was able to complete them. These challenges also affected older adults’ self-efficacy, with P7 saying, ‘I just think I’m not smart enough to do it’; ‘I just didn’t understand anything they were telling me to do... I’m a complete failure’; and ‘I don’t even understand the instructions. Is everybody else that does this as dumb as I am?’ (Brewer, Morris and Piper, 2016)
“Reflexivity”

- The intuitions of Social Scientists are not merely insufficient; they’re *actively misleading*
“Reflexivity”

- The intuitions of Social Scientists are not merely insufficient; they’re *actively misleading*.

- Descriptive (often qualitative) research is necessary to motivate research questions.
Considerable energy has been spent investigating the phenomenon of echo chambers.
“Reflexivity”: Online Echo Chambers

- Considerable energy has been spent investigating the phenomenon of echo chambers.

- They don’t exist...except among users in specialized (partisan or professional) networks.
“Reflexivity”: Online Echo Chambers

- Considerable energy has been spent investigating the phenomenon of echo chambers.

- They don’t exist...except among users in specialized (partisan or professional) networks.
  - Including, of course, Social Scientists and journalists.
“Reflexivity”: Online Echo Chambers

- Considerable energy has been spent investigating the phenomenon of echo chambers

- They don’t exist...except among users in specialized (partisan or professional) networks
  - Including, of course, Social Scientists and journalists

- Science works: empirical consensus falsified the theory of ubiquitous echo chambers
“Reflexivity”: Online Echo Chambers

- Considerable energy has been spent investigating the phenomenon of echo chambers.

- They don’t exist...except among users in specialized (partisan or professional) networks.
  - Including, of course, Social Scientists and journalists.

- Science works: empirical consensus falsified the theory of ubiquitous echo chambers.

- But the supply of social science research is inelastic, so there are serious opportunity costs.
## Opportunity Costs

<table>
<thead>
<tr>
<th></th>
<th>Online Echo Chambers</th>
<th>Fake News</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced by academics</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

TOO MUCH TOO LATE
## Opportunity Costs

<table>
<thead>
<tr>
<th></th>
<th>Online Echo Chambers</th>
<th>Fake News</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced by academics</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Experienced by public</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>
Opportunity Costs

<table>
<thead>
<tr>
<th></th>
<th>Online Echo Chambers</th>
<th>Fake News</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced by academics</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Experienced by public</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Specific sub-populations</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>
## Opportunity Costs

<table>
<thead>
<tr>
<th></th>
<th>Online Echo Chambers</th>
<th>Fake News</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced by academics</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Experienced by public</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Specific sub-populations</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Studied by academics1</td>
<td>TOO MUCH</td>
<td>TOO LATE</td>
</tr>
</tbody>
</table>
Clickbait Effects on the Clickers

“The Generalizability of Survey Experiments” (Mullinix et al., 2015):

- Some convenience samples would be inappropriate such as a student sample where a moderator is age.
- Any sample with a hard digital literacy cutoff is inappropriate for making generalizations about online behaviors.
- Clearly excludes MTurk.
- Excludes even our over-sample of digital naives due to difficulty of survey instruments we never thought existed.
Clickbait Effects on the Clickers

“"The Generalizability of Survey Experiments" (Mullinix et al., 2015):

- General point: survey experiments generalize...

- Some convenience samples would be inappropriate such as a student sample where a moderator is age.

- Any sample with a hard digital literacy cutoff is inappropriate for making generalizations about online behaviors.

- Clearly excludes MTurk.

- Excludes even our over-sample of digital naives due to difficulty of survey instruments we never thought existed.
Clickbait Effects on the Clickers

“The Generalizability of Survey Experiments” (Mullinix et al., 2015):

- General point: survey experiments generalize...
  - “some convenience samples would be inappropriate such as a student sample where a moderator is age”
Clickbait Effects on the Clickers

- “The Generalizability of Survey Experiments” (Mullinix et al., 2015):
  - General point: survey experiments generalize...
  - “some convenience samples would be inappropriate such as a student sample where a moderator is age”

- Any sample with a hard digital literacy cutoff is inappropriate for making generalizations about online behaviors
“The Generalizability of Survey Experiments” (Mullinix et al., 2015):

- General point: survey experiments generalize...
- “some convenience samples would be inappropriate such as a student sample where a moderator is age”

Any sample with a hard digital literacy cutoff is inappropriate for making generalizations about online behaviors

Clearly excludes MTurk
Clickbait Effects on the Clickers

“The Generalizability of Survey Experiments” (Mullinix et al., 2015):
- General point: survey experiments generalize...
- “some convenience samples would be inappropriate such as a student sample where a moderator is age”

- Any sample with a hard digital literacy cutoff is inappropriate for making generalizations about online behaviors

- Clearly excludes MTurk

- Excludes even our over-sample of digital naives due to difficulty of survey instruments we never thought existed
Thank You
