

# **Item Count Techniques** *2.0*

Julian Jamison

University of Exeter &  
The World Bank

Measuring the Tricky things  
Oxford: October 12, 2018

# Motivation

- Presumably don't need much here, but...
- How can we get reliable data regarding
  - violent criminal behavior?
  - risky sexual behavior?
  - bias against minority groups?
  - gender-based violence?
  - etc.

# Motivation

- For my purpose here, what these have in common is a potential concern with respect to *privacy*, *shame*, fear of *consequences*, and so on

# Motivation

- For my purpose here, what these have in common is a potential concern with respect to *privacy*, *shame*, fear of *consequences*, and so on
- This technique will not be as useful when the concern is
  - respondent unaware of truth (e.g. profits)
  - sensitive... but internalized stigma and/or self-deception
  - desire to influence results / conclusions

# Alternatives

- Randomized response techniques
  - Coin flip with forced “yes” to sensitive question (may be difficult to admit, even if believe masked)
  - Coin flip to decide which of two questions to answer
  - Die roll for cardinal queries

# Alternatives

- Randomized response techniques
  - Coin flip with forced “yes” to sensitive question (may be difficult to admit, even if believe masked)
  - Coin flip to decide which of two questions to answer
  - Die roll for cardinal queries
- Qualitative ‘ground-truthing’

# Alternatives

- Randomized response techniques
  - Coin flip with forced “yes” to sensitive question (may be difficult to admit, even if believe masked)
  - Coin flip to decide which of two questions to answer
  - Die roll for cardinal queries
- Qualitative ‘ground-truthing’
- And of course many more, including careful survey design and deflection to e.g. “others like you”

# Item Count Technique

- Mask or veil the individual-level response by mixing the sensitive behavior or attitude with a number of other *innocuous* statements
- Aka unmatched count or list response



# Item Count Technique

- Mask or veil the individual-level response by mixing the sensitive behavior or attitude with a number of other *innocuous* statements
- Aka unmatched count or list response
- Dates to Raghavarao & Federer (1979)
- Used in psychology (Dalton *et al.* 1994)
- and political science (e.g. Corstange 2009)

# Example from Uganda

129.	<p><b>ONLY USE THIS IF LIST 1 = "A"</b></p> <p>Please tell me how many of the following statements are true for you, but not which ones in particular:</p> <ol style="list-style-type: none"><li>1. Your family owns a house</li><li>2. Your biological father is alive</li><li>3. You raise goats</li><li>4. You really like posho</li></ol> <p><b>WRITE 9 IF REFUSED TO ANSWER</b></p>	<p>NUMBER OF TRUE STATEMENTS.....</p> <input data-bbox="1746 596 1839 668" type="text"/>
130.	<p><b>ONLY USE THIS IF LIST 1 = "B"</b></p> <p>Please tell me how many of the following statements are true for you, but not which ones in particular:</p> <ol style="list-style-type: none"><li>1. Your family owns a house</li><li>2. Your biological father is alive</li><li>3. You raise goats</li><li>4. You really like posho</li><li>5. The <b>last time</b> you had sexual intercourse, a male condom was used.</li></ol> <p><b>WRITE 9 IF REFUSED TO ANSWER</b></p>	<p>NUMBER OF TRUE STATEMENTS.....</p> <input data-bbox="1746 1162 1839 1233" type="text"/>

# Immediate issues

- Responses are not masked at boundaries
  - Can minimize chance of this happening

# Immediate issues

- Responses are not masked at boundaries
  - Can minimize chance of this happening
- Use precious sample size to estimate distribution of innocuous questions
  - Can be a relatively small fraction (e.g. 10%)
  - or even out-of-sample from same population

# Immediate issues

- Responses are not masked at boundaries
  - Can minimize chance of this happening
- Use precious sample size to estimate distribution of innocuous questions
  - Can be a relatively small fraction (e.g. 10%)
  - or even out-of-sample from same population
- Only aggregate data for analysis
  - No: still individual level, just noisier (and not binary)
  - Can even refine on the basis of observables

# Results: infidelity

Ever unfaithful in the past three months? (%)

	Direct	Indirect	p-value*
Overall	13.3	18.9	0.04
Male	19.6	21.9	big
Female	7.1	16.3	0.02

\*one-sided

# Results: safe sex

Used a condom the last time you had sex? (%)

	Direct	Indirect	p-value*
Overall	24.3	24.0	really big
Male	29.6	22.5	0.06
Female	19.4	25.4	0.09
Single	45.5	32.4	0.01
Partnered	11.7	18.7	0.04

\*one-sided

NB: female results driven entirely by partnered

# Choosing the innocuous questions

- Should their topic matter be
  - similar to the sensitive question? (hiding technique...)
  - totally different? (increasing transparency – and saliency)



# Choosing the innocuous questions

- Should their topic matter be
  - similar to the sensitive question? (hiding technique...)
  - totally different? (increasing transparency – and saliency)
- No right answer, but my default is the latter
  - The whole point is to be clear how it is masked
  - More freedom to fine-tune, as on next slide

# Choosing the innocuous questions

- Should their topic matter be
  - similar to the sensitive question? (hiding technique...)
  - totally different? (increasing transparency – and saliency)
- No right answer, but my default is the latter
  - The whole point is to be clear how it is masked
  - More freedom to fine-tune, as on next slide
- But: maybe not if salience could be a problem
  - E.g. unwilling to self-acknowledge; desire to distort data

# Choosing the innocuous questions

- Optimization of relevant properties
- Tradeoff regarding how many:
  - Too few doesn't mask as well, hits boundaries
  - Too many takes time and increases errors

# Choosing the innocuous questions

- Optimization of relevant properties
- Tradeoff regarding how many:
  - Too few doesn't mask as well, hits boundaries
  - Too many takes time and increases errors
- Ideally 50% "yes" for each
  - If e.g. two negatively correlated pairs (Glynn 2013), then reduce probability of boundary and also noise

# Choosing the innocuous questions

- Where should sensitive question appear in order?

# Choosing the innocuous questions

- Where should sensitive question appear in order?
- Does the framing of sensitive question matter?
  - Indeed: better to have the sensitive response be “no”

# Choosing the innocuous questions

- Where should sensitive question appear in order?
- Does the framing of sensitive question matter?
  - Indeed: better to have the sensitive response be “no”
- Can we validate the technique?
  - Tsuchiya *et al.* (2007) do a placebo test
  - Blair *et al.* (2014) confirm similar results as ‘endorsement experiment’ which is quite distinct
  - Generally at least compare direction of effect to theory

# Coffman *et al.* (2017)

- Roughly state-of-the-art, putting together many of the points discussed
- “ever had same-sex sexual experience”
  - 17.2% direct, goes up to 27.4% if veiled
- “should be illegal to discriminate against LGBT”
  - 85.6% direct, goes down to 75.3% if veiled



# Unresolved issues: consider GBV

- Ask the [possible] perpetrators, or the [possible] victims / survivors? Legal requirements?

# Unresolved issues: consider GBV

- Ask the [possible] perpetrators, or the [possible] victims / survivors? Legal requirements?
- Role of technology, e.g. respondent holds tablet privately; items disappear off the screen; etc.
  - though NB Coffman paper was anonymous online sample

# Unresolved issues: consider GBV

- Ask the [possible] perpetrators, or the [possible] victims / survivors? Legal requirements?
- Role of technology, e.g. respondent holds tablet privately; items disappear off the screen; etc.
  - though NB Coffman paper was anonymous online sample
- ICT may require higher sample size (cost), but avoids cost of training enumerators and of referrals

# Unresolved issues: consider GBV

- Ask the [possible] perpetrators, or the [possible] victims / survivors? Legal requirements?
- Role of technology, e.g. respondent holds tablet privately; items disappear off the screen; etc.
  - though NB Coffman paper was anonymous online sample
- ICT may require higher sample size (cost), but avoids cost of training enumerators and of referrals
- Is it ethical to purposely avoid e.g. referrals? Is it ethical not to get realistic measure of what works?

# 'Meta'-level masking

- Poor data in the US on gun ownership, and even more so fraction who have used (fired or threatened) a gun defending self / family / home
- Difficulty with the latter is that a subgroup may have incentive to exaggerate *so that* the results look bigger, for policy purposes

# 'Meta'-level masking

- Poor data in the US on gun ownership, and even more so fraction who have used (fired or threatened) a gun defending self / family / home
- Difficulty with the latter is that a subgroup may have incentive to exaggerate *so that* the results look bigger, for policy purposes
- One idea is to prime them to think research is on voting, crime, or sexual behavior but then actually have the target question be as above...

# Conclusion

- There have also been criticisms of the method, including instances where it does not seem to have worked (as expected)
- Surely imperfect (e.g. usually only better, not actually right) but on the whole seems to be effective and useful

# Conclusion

- There have also been criticisms of the method, including instances where it does not seem to have worked (as expected)
- Surely imperfect (e.g. usually only better, not actually right) but on the whole seems to be effective and useful
- IMO tradeoffs are more around cost (time, money) and applicability to subset of sensitive topics
- Fortunately this gives us all something to do...