Monitoring Harassment in Organizations

Laura Boudreau[1] Sylvain Chassang[2] Ada González-Torres[3] Rachel Heath[4]

[1]Columbia University

[2]Princeton University

[3]Ben Gurion University

[4]University of Washington

December 6, 2024

Motivation

- Organizations' ability to act against harassment is limited by their ability to collect information from relevant parties.
  - Victims and witnesses concerned about possible retaliation and reputational costs.

Motivation

- Organizations' ability to act against harassment is limited by their ability to collect information from relevant parties.
  - Victims and witnesses concerned about possible retaliation and reputational costs.

- Prevents organizations from assessing the scope and nature of harassment and from acting on the problem.
  - How prevalent is harassment?
  - What share of employees is responsible for the damage?
  - How isolated are victims across teams?

Motivation

- Organizations' ability to act against harassment is limited by their ability to collect information from relevant parties.
  - Victims and witnesses concerned about possible retaliation and reputational costs.

- Prevents organizations from assessing the scope and nature of harassment and from acting on the problem.
  - How prevalent is harassment?
  - What share of employees is responsible for the damage?
  - How isolated are victims across teams?

- Theory predicts that providing **plausible deniability** through **hard garbling** can improve information transmission.
  - Randomly record a fixed share of reported "No's" as "Yes's".
  - Warner, 1965; Chassang & Padró i Miquel, 2018; Chassang and Zehnder, 2019

This research

- Test impact of survey design on reporting misbehavior:
  1. Hard garbling
     - Contrast to soft garbling/randomized response.
  2. Rapport building
  3. Removing team-level information

- Using improved survey data, assess policy-relevant aspects of harassment.

Bangladeshi producer concerned about harassment

Source: 47roots.com.

- Partner with senior management of large apparel producer that aims to use workers' feedback to improve its HR policies.
- Conduct phone-based survey experiment about harassment with 2,245 workers at 2 plants.
- Randomize survey method (direct or garbled), degree of rapport built, and degree of team-level info (manager id or not).

Anonymous apparel producer partner

- Among largest in Bangladesh.
- $\approx$ 80% of workers female, but 90%+ of managers male.
- Performs poorly on workers' job satisfaction and lower-level managers' performance relative to competitors.
- Senior management concerned that it may only be aware of the "tip of the iceberg."
  - Aimed to improve its awareness of nature and scope of managers' mistreatment of workers.
  - Longer-term goal: Collect continuous feedback from workers and tie to managers' incentives.

Outline

Set-up for experiment

- Phone-based survey with employees at two factories that produce denim products.
- Conducted by BRAC Institute of Governance and Development (BIGD).
- Participants informed that results would be shared with senior management and would inform HR policy.
- Pre-analysis plan pre-registered on AEA RCT Registry.

Outcomes: Experience of harassment

In the past year, has your line supervisor taken any of the following actions toward you against your will? (Yes/No)

Mistreatment	Actions read aloud to respondent:
Threatening behavior	Threatened you; Told you that they will harm you if you do not agree to or fulfill their demands.
Physical harassment	Hit, slapped, or punched you; Cut or stabbed you; Tripped you; Otherwise intentionally caused you physical harm.
Sexual harassment	Made remarks about you in a sexual manner; Asked you to enter into a love or sexual relationship; Asked or forced you to perform sexual favors; Asked or forced you to meet outside of the factory or meet them alone in a way that made you feel uncomfortable; Touched you in a sexual manner or in a way that made you feel uncomfortable or scared; Shown you pictures of sexual activities.

Treatment conditions

- Survey method for harassment-related questions Scripts:

  1.a) Direct Elicitiation (DE)

  1.b) Hard Garbling (HG) with $\pi = 0.2$ ex ante.

Treatment conditions

- Survey method for harassment-related questions Scripts:

  1.a) Direct Elicitiation (DE)

  1.b) Hard Garbling (HG) with $\pi = 0.2$ ex ante.

- Personally-identifying information (PII):

  2.a) Status quo: Ask questions that reveal relatively more PII: production section/line number and direct supervisor.

  2.b) Low PII: Limit PII requested; no questions asked about production section/line number or direct supervisor.

Treatment conditions

- Survey method for harassment-related questions Scripts:

  1.a) Direct Elicitiation (DE)

  1.b) Hard Garbling (HG) with $\pi = 0.2$ ex ante.

- Personally-identifying information (PII):

  2.a) Status quo: Ask questions that reveal relatively more PII: production section/line number and direct supervisor.

  2.b) Low PII: Limit PII requested; no questions asked about production section/line number or direct supervisor.

Note: Trade-off: Cannot calculate manager-level statistics.

Treatment conditions

- Survey method for harassment-related questions Scripts:
  - 1.a) Direct Elicitiation (DE)
  - 1.b) Hard Garbling (HG) with $\pi = 0.2$ ex ante.
- Personally-identifying information (PII):
  - 2.a) Status quo: Ask questions that reveal relatively more PII: production section/line number and direct supervisor.
  - 2.b) Low PII: Limit PII requested; no questions asked about production section/line number or direct supervisor.
- Rapport-building (RB):
  - 3.a) Status quo: Typical social science intro script and survey.
  - 3.b) RB: Survey enumerators allocate survey time to build rapport, or trust, with the participant. Short v long RB

Summary statistics (N=2144)

	Mean	SD	Min	p25	p50	p75	Max
Female	0.81	0.39	0	1	1	1	1
Currently Working	0.96	0.20	0	1	1	1	1
Age	26.8	5.14	17	23	26	30	55
Experience (yrs)	5.20	3.57	0	2.83	4.42	7.17	28.8
Tenure (yrs)	2.91	2.45	0.0027	0.61	2.87	4.21	16.9
Tenure in Team (yrs) [N=1554]	2.58	2.52	0	0.50	1.92	3.92	14.5
Years of Education	6.71	3.38	0	5	6.50	9	16
Marital Status (1=Yes)	0.82	0.38	0	1	1	1	1
Children (1=Yes)	0.74	0.44	0	0	1	1	1
Sewing Section	0.49	0.50	0	0	0	1	1
Finishing Section	0.34	0.47	0	0	0	1	1
Washing Section	0.17	0.38	0	0	0	0	1
Position: Helper	0.18	0.39	0	0	0	0	1
Position: Ironing/Folding	0.086	0.28	0	0	0	0	1
Position: Operator	0.59	0.49	0	0	1	1	1
Position: Packer	0.046	0.21	0	0	0	0	1
Position: Quality	0.096	0.29	0	0	0	0	1

Main regression model

$$\hat{r}_{is} = \alpha\, HG_i + \beta\, LowPII_i + \gamma\, Rapport_i + \mu_s + \theta X_i + \xi_{is} \qquad (1)$$

- $\hat{r}_{is}$: Transformed reporting outcome for individual i in stratum s.
- $HG_i, LowPII_i$ and $Rapport_i$: Hard-garbling, asking for less PII, and rapport building, respectively.
- μ_s: Stratum fixed-effects.
- X_i: Controls for individuals' characteristics, selected using PDS lasso (Belloni, Chernozhukov, and Hansen, 2014).
- ξ_{is}: residual term; robust standard errors reported.

Outline

Impacts of survey design

	Threatening behavior		Physical harassment		Sexual harassment	
	(1)	(2)	(3)	(4)	(5)	(6)
HG	0.0453**	0.0453**	0.0437**	0.0437**	0.0478**	0.0478**
	(0.0213)	(0.0206)	(0.0189)	(0.0182)	(0.0193)	(0.0186)
Rapport	0.0128	0.0128	-0.0094	-0.0094	0.0188	0.0188
	(0.0200)	(0.0194)	(0.0176)	(0.0170)	(0.0182)	(0.0176)
Low PII	0.0092	0.0092	0.0280	0.0280	0.0045	0.0045
	(0.0227)	(0.0219)	(0.0187)	(0.0181)	(0.0190)	(0.0184)
Control Mean	.099	.099	.0152	.0152	.0178	.0178
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
PDS lasso controls	No	Yes	No	Yes	No	Yes
Observations	2141	2141	2141	2141	2141	2141

Impacts of survey design: Heterogeneity by gender

	Threatening behavior		Physical harassment		Sexual harassment	
	(1)	(2)	(3)	(4)	(5)	(6)
HG Treatment × Female	0.0283	0.0269	0.0405*	0.0414**	0.0365*	0.0385*
	(0.0234)	(0.0226)	(0.0209)	(0.0201)	(0.0214)	(0.0206)
HG Treatment × Male	0.1230**	0.1216**	0.0597	0.0585	0.0938**	0.0930**
	(0.0514)	(0.0498)	(0.0452)	(0.0437)	(0.0452)	(0.0440)
Rapport × Female	0.0215	0.0229	-0.0173	-0.0162	0.0336*	0.0339*
	(0.0222)	(0.0214)	(0.0195)	(0.0187)	(0.0202)	(0.0194)
Rapport × Male	-0.0250	-0.0250	0.0243	0.0193	-0.0512	-0.0510
	(0.0472)	(0.0454)	(0.0416)	(0.0403)	(0.0423)	(0.0413)
Low PII Treatment × Female	0.0119	0.0109	0.0325	0.0324	0.0097	0.0107
	(0.0252)	(0.0242)	(0.0207)	(0.0198)	(0.0215)	(0.0207)
Low PII Treatment × Male	-0.0067	-0.0089	0.0120	0.0169	-0.0237	-0.0211
	(0.0535)	(0.0517)	(0.0443)	(0.0432)	(0.0417)	(0.0404)
Control Mean - Female	.0798	.0798	.0092	.0092	.0184	.0184
Control Mean - Male	.1912	.1912	.0441	.0441	.0147	.0147
p(HGxFemale - HGxMale)	[0.094]	[0.084]	[0.700]	[0.724]	[0.253]	[0.264]
p(RapportxFemale - RapportxMale)	[0.373]	[0.340]	[0.366]	[0.425]	[0.071]	[0.063]
p(NoPIIxFemale - NoPIIxMale)	[0.753]	[0.729]	[0.675]	[0.745]	[0.478]	[0.487]
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
PDS lasso controls	No	Lasso	No	Lasso	No	Lasso
Observations	2141	2141	2141	2141	2141	2141

Outline

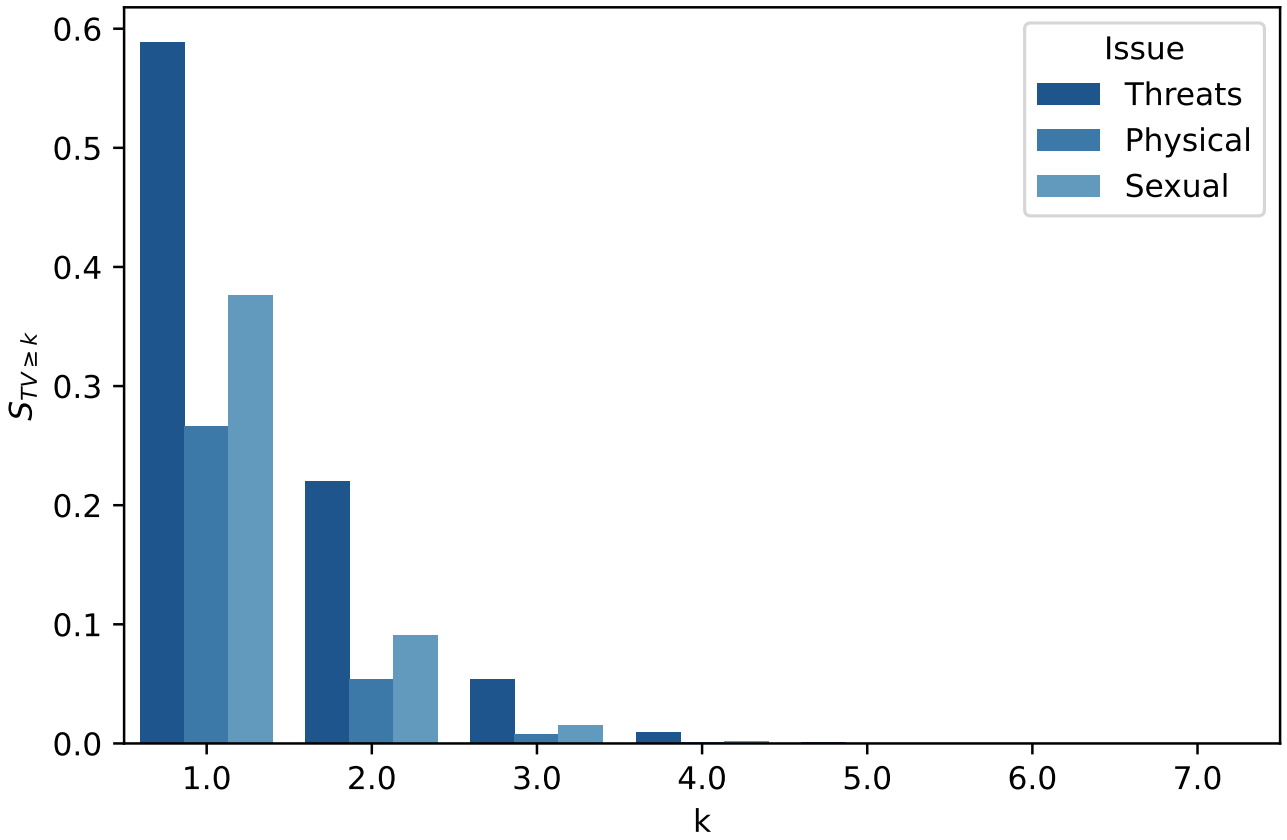
Harassment underdetected (DE), yet widespread (HG)

Share of workers who have been victimized:

Notes: We pool across all treatment arms, including the RB arms and the arms in which we do not collect team-level identifying information.

Most managers harass

Median of 7 workers per team in HG/PII arms (112 teams). For teams of size 7, share of teams with at least k victims:



Discussion

- Lack of plausible deniability causes severe under-reporting of harassment in an organizational setting.

Discussion

- Lack of plausible deniability causes severe under-reporting of harassment in an organizational setting.
- Rapport building helps for women, but may backfire for men.

Discussion

- Lack of plausible deniability causes severe under-reporting of harassment in an organizational setting.
- Rapport building helps for women, but may backfire for men.
- Removing team-level identifying information does not help and comes at cost of being unable to compute manager-level statistics.

Discussion

- Lack of plausible deniability causes severe under-reporting of harassment in an organizational setting.
- Rapport building helps for women, but may backfire for men.
- Removing team-level identifying information does not help and comes at cost of being unable to compute manager-level statistics.
- Hard garbling dramatically increases reporting, reducing measurement error in key statistics.
  - Promising results with broad relevance, but need more testing.

Discussion

- Lack of plausible deniability causes severe under-reporting of harassment in an organizational setting.
- Rapport building helps for women, but may backfire for men.
- Removing team-level identifying information does not help and comes at cost of being unable to compute manager-level statistics.
- Hard garbling dramatically increases reporting, reducing measurement error in key statistics.
- Directions for future research:
  - How to scale up enforcement actions taken as a function of reports?
  - Welfare/distributional implications of improved governance systems for workers, managers, and producers?

Short- versus long-run effects

In long run:

- Managers may increase magnitude of retaliation in response to increased anonymity provided by HG;
- Workers may start strategically misreporting well-behaving managers.

  Makes it difficult to interpret a drop or an increase in incriminating reports.

Chassang and Padró i Miquel (2018) show that can reach whistleblowing policies that deliver robust guarantees on underlying level of misbehavior:

- Starting from low level of enforcement, reduce info content of reports up to a point where workers are willing to complain;
- Keeping info content of positive reports $(Pr(r = 1|\tilde{r} = 1)/Pr(r = 0|\tilde{r} = 1))$ the same, scale up enforcement.

Measurement using garbled data (cont.)

- Denote the sample distribution of profiles of intended and recorded reports across teams with $\hat{\mu} \in \Delta(\{0,1\}^I)$ and $\hat{\mu} \in \Delta(\{0,1\}^I)$, respectively:

$$\forall r \in \{0,1\}^I, \quad \hat{\mu}(r) \equiv \frac{1}{m} \sum_{a \in M} 1_{r_a = r} \quad \text{and} \quad \tilde{\mu}(r) \equiv \frac{1}{m} \sum_{a \in M} 1_{\tilde{r}_a = r}.$$

Proposition 2 (Identification from garbled reports)

As m grows large, the sample distribution of intended reports $\hat{\mu}$ is identified from the sample distribution of recorded reports $\hat{\mu}$.

Proof

◂ Measurement

DE Script

We are now going to ask you several questions about the way your manager treats you and other employees.

For instance: Has your manager shouted at you in the last month? Yes or No?

Each of the questions has a Yes or No answer. Your answers will be recorded as you go, but we can chat about them before we record them for good.

HG Script

We are now going to ask you several questions about the way your manager treats you and other employees. For instance: Has your manager shouted at you in the last month? Yes or No?

Each of the questions has a Yes or No answer. Our system is setup so that it's safe to report an issue. If you choose to respond YES (there is an issue), our system will

record it as a YES for sure.

Importantly, if someone responds NO, the system will sometimes record the response as YES.

This means that if you respond YES, we can guarantee that you won't be the only person saying YES. For every 5 responses from workers, at least 1 will be recorded as YES.

The researchers are only interested in the total number of yes/no responses from all surveys. If you respond YES, aside from me, no one will ever be able to know that this was your answer, not even the researchers. Your answers are fully protected with us.

◂ Treatment conditions

HG Script: Comprehension questions

Before we begin the survey questions, we would like to check whether we have explained our survey system clearly. Can you please tell me whether the following statements are TRUE or FALSE.

a. If I respond “Yes,” no one can ever know this for sure.

b. The system will record at least one out of every five workers’ responses as “Yes.”

Instructions to survey enumerator: Survey enumerator reports correct answers to respondent after asking both questions:

“It is true that if you respond “Yes,” the system is designed so that no one can ever know this for sure. And it is also true that the system will record at least one out of every five workers’ responses as yes, so we can guarantee that anyone who says yes will not be the only person saying yes.”

Note that we can chat about your answers before we record them for good, but I don’t know whether the system would record a NO as a YES.

Short versus long RB conditions

- RB1: Enumerator signals care using emotional mirroring and acknowledgment.
- RB2: Extended RB section, worker has chance to ask enumerator questions. Enumerator shares a related experience.

Motivating evidence on reporting

- 2017 survey of garment workers recruited through the community, conducted by BRAC Institute of Governance and Development (BIGD).

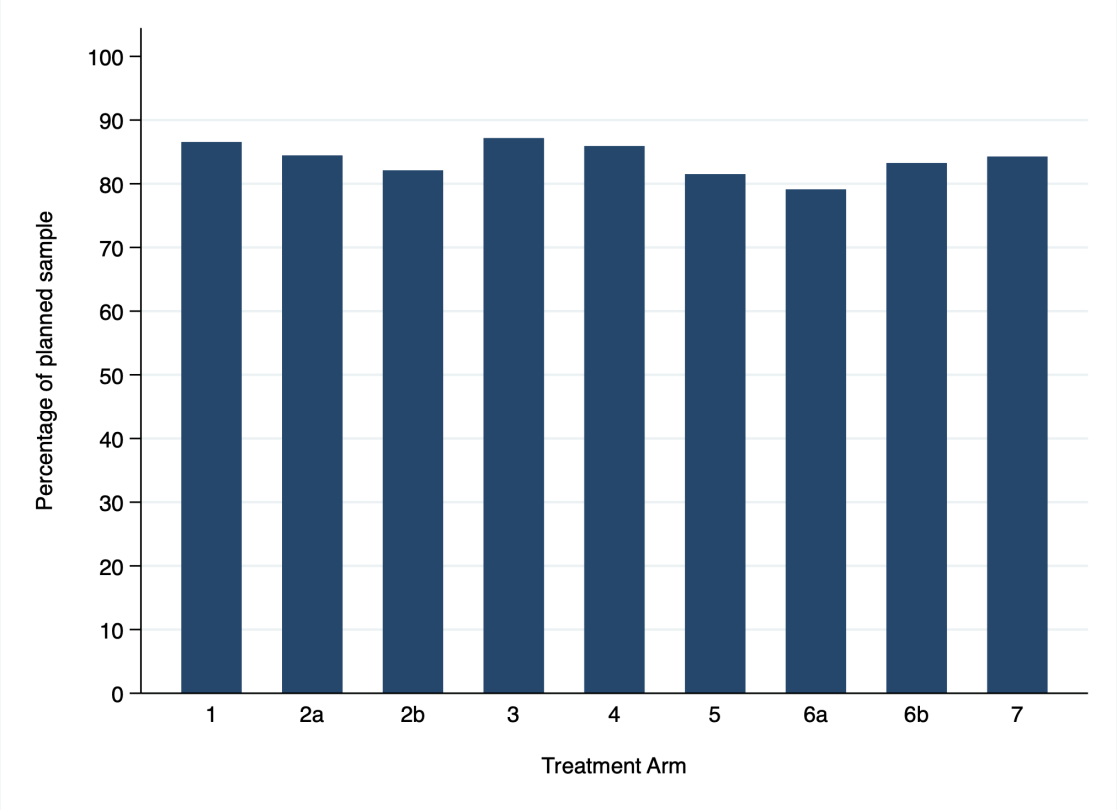
Table: Workers' reported experience of harassment (N=1500)

Variable	Proportion who respond "Yes"
Witnessed physical harassment	0.201
Experienced physical harassment	0.011
Witnessed sexual harassment	0.111
Experienced sexual harassment	0.001

Source: Authors' calculations using data from Kabeer, Huq, and Sulaiman (2020).

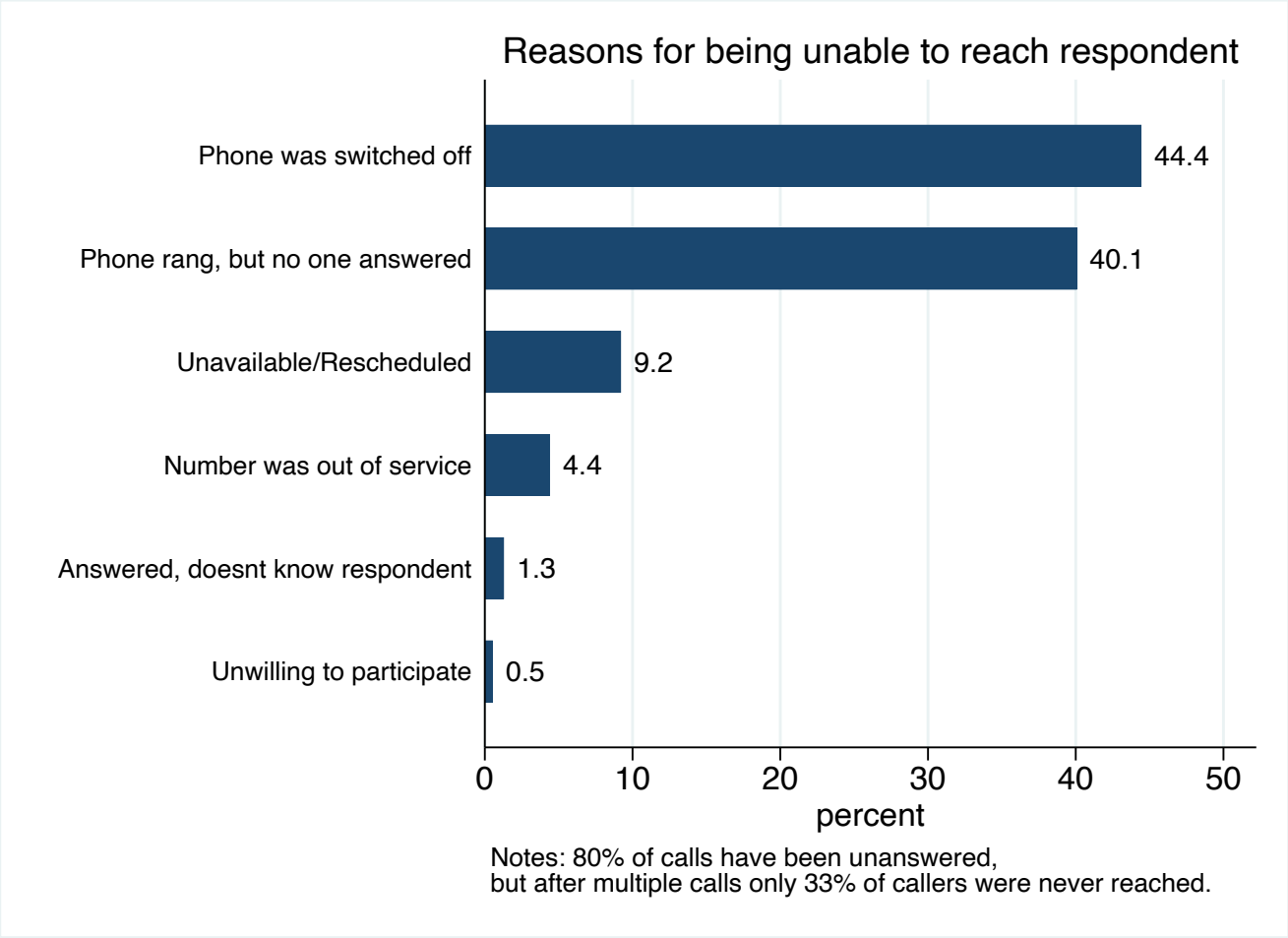
◂ Reporting & Garbling ◂ back to Context

Response rate balance across treatment arms



◂ Sample & Randomization

Reasons for non-response



◂ Sample & Randomization