How informed is consent?

A field experiment

Alexandra Avdeenko and Matthias Stelter

Research question

- Increasingly more primary data collections and field research in low- and middle-income countries
- Increasing awareness of data protection and ownership issues
- Research ethics and laws dictate obtaining informed consent of study participants
- How informed is consent? Can we increase informedness with alternative approaches?

$\Rightarrow$ Measure understanding and experimentally test augmented approaches for obtaining informed consent in an large scale data collection in Pakistan

Related literature I

Research ethics:

- Push for more attention on research ethics (e.g., Asiedu, Karlan, Lambon-Quayefio, and Udry (2021))
- Ethics of informed consent (e.g., Cassell and Young (2002), Hutton, Eccles, and Grimshaw (2008))

Data collection features can affect sample composition:

- Financial incentives affect response rates and sample composition in surveys (Singer and Ye (2013) and Mercer, Caporaso, Cantor, and Townsend (2015))
- Requirement for consent can threaten research studies (Tu, Willison, Silver, Fang, Richards, Laupacis, and Kapral (2004))

Related literature II

Understanding of consent:

- Consumers are unaware of privacy threats from sharing personal information (Acquisti, Taylor, and Wagman (2016))
- Patients have a poor understanding of the risks and complications of their medical procedures (e.g., Falagas, Korbila, Giannopoulou, Kondilis, and Peppas (2009); Stanley, Walters, and Maddern (1998))
- What processes can improve understanding in medical research participants? (e.g., Fitzgerald, Marotte, Verdier, Johnson, and Pape (2002) and Stunkel, Benson, McLellan, Sinaii, Bedarida, Emanuel, and Grady (2010))

Contribution:
First empirical evidence for survey data collection in low- and middle-income countries

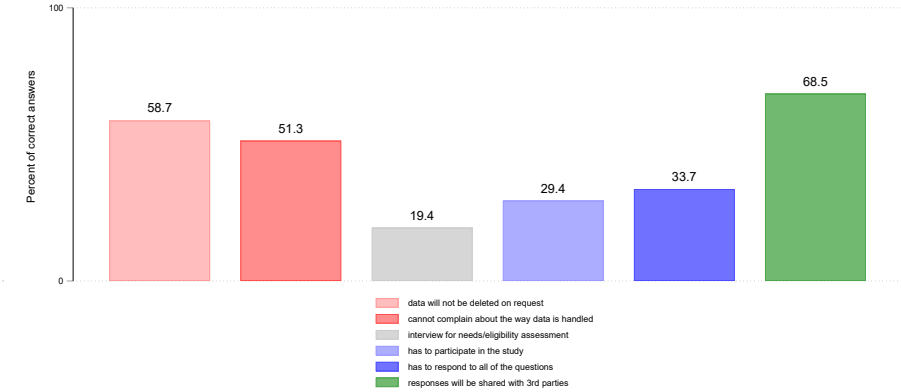
Sample and context

Two large data collections: Sindh (winter 2020/2021) and Punjab (summer 2021), Pakistan

	Control		Video		+Dialogue		Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	mean	sd	mean	sd	mean	sd	V-C	D-C	D-V
Female	0.61	0.49	0.63	0.48	0.58	0.49	0.01	-0.04	-0.05**
Age	42.1	11.6	42.1	11.3	42.5	11.9	-0.03	0.46	0.49
Household head	0.37	0.48	0.36	0.48	0.40	0.49	-0.00	0.03	0.04
Spouse	0.49	0.50	0.50	0.50	0.48	0.50	0.01	-0.02	-0.02
Other relation	0.14	0.35	0.14	0.34	0.12	0.33	-0.01	-0.02	-0.01
Number of observations:									
Meta indicators	3903		3404		445		7307	4348	3849
Objective item level	1003		559		444		1562	1447	1003
Subjective item level	1312		856		444		2168	1756	1300
Full module	446		0		445		0	891	0

$* : p < 0.10;\ \ ** : p < 0.05;\ \ *** : p < 0.01$

Finding Ia: Limited understanding



Note: Share of correct responses when false is correct in benchmark group.

Figure: Objective measure: False correct

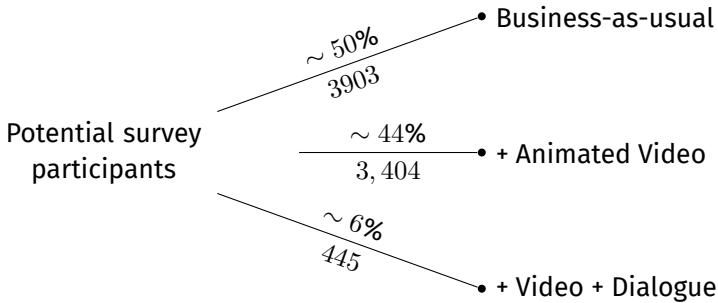
True correct version | Implementation | Enumerator pilot

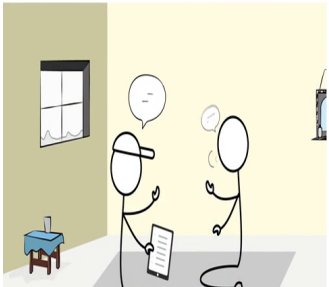
Finding Ib: Limited understanding

Note: Share of respondents that reported to understand respective aspects well or fully in benchmark group.

Figure: Subjective measure

The experiment





Estimation equation:

$$Y = \alpha + \beta_1 D_{Video} + \beta_2 D_{+Dialogue} + \xi + u$$

Finding II: Consent rate

	Consent Rate		Response Rate
	(1) 1st visit	(2) 2nd visit	(3) 2nd visit
Video (β_1)	0.001	-0.000	0.003
	(0.001)	(0.001)	(0.004)
+Dialogue (β_2)	0.001	0.000	0.012*
	(0.002)	(0.000)	(0.007)
Diff. (β_2-β_1)	-0.00	0.00	0.01
	(0.00)	(0.00)	(0.01)
Model description:			
Interviewer FE	Yes	Yes	Yes
Adj. R^2	0.03	0.01	0.04
Control group mean	1.00	1.00	0.97
Observations	7752	7309	7736

$* : p < 0.10; \quad ** : p < 0.05; \quad *** : p < 0.01$

Findings III: Objective measure of understanding

	Overall		Rights		Purpose	Voluntariness		Confidentiality
	(1) Score	(2) Informed	(3) Delete info.	(4) Can compl.	(5)	(6) Part. obliged	(7) Resp. obliged	(8)
Video (β_1)			-0.014	-0.111***	-0.028	-0.025	-0.003	-0.063**
			(0.030)	(0.031)	(0.020)	(0.024)	(0.025)	(0.029)
+Dialogue (β_2)	0.011	0.033	0.004	-0.031	0.001	0.048*	0.106***	-0.021
	(0.012)	(0.026)	(0.037)	(0.034)	(0.027)	(0.027)	(0.031)	(0.034)
Diff. (β_2-β_1)			0.02	0.08**	0.03	0.07**	0.11***	0.04
			(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Model description:								
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.58	0.52	0.48	0.53	0.45	0.50	0.46	0.51
Control group mean	0.67	0.38	0.59	0.51	0.19	0.29	0.34	0.69
Observations	889	889	807	820	1216	1225	1207	810

$* : p < 0.10; \quad ** : p < 0.05; \quad *** : p < 0.01$

Heterogenous effects

Findings IV: Subjective measure of understanding

	Overall			Rights	Purpose	Voluntariness	Confidentiality
	(1) Share	(2) All	(3) Score	(4)	(5)	(6)	(7)
Video (β_1)				0.002	-0.007	-0.022	0.004
				(0.017)	(0.016)	(0.016)	(0.016)
+Dialogue (β_2)	-0.042**	-0.048*	-0.080**	-0.023	-0.060***	-0.054***	-0.032
	(0.020)	(0.025)	(0.035)	(0.021)	(0.021)	(0.021)	(0.022)
Diff. (β_2-β_1)				-0.03	-0.05**	-0.03	-0.04
				(0.02)	(0.02)	(0.02)	(0.02)
Model description:							
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.66	0.59	0.64	0.47	0.47	0.49	0.48
Control group mean	0.59	0.41	3.56	0.57	0.61	0.64	0.61
Observations	882	882	882	2561	2612	2606	2616

$* : p < 0.10; \quad ** : p < 0.05; \quad *** : p < 0.01$

Subjective vs. objective

Findings V: Item non-response rates

	1st Visit		2nd Visit (HH Roster)		2nd Visit (Full Int.)	
	(1) Non-resp. rate	(2) Any non-resp.	(3) Non-resp. rate	(4) Any non-resp.	(5) Non-resp. rate	(6) Any non-resp.
Video (β_1)	-0.038	-0.003	0.004	0.001	0.009	0.010
	(0.045)	(0.003)	(0.010)	(0.005)	(0.024)	(0.011)
+Dialogue (β_2)	0.034	0.002	-0.022	-0.008	-0.056*	0.005
	(0.091)	(0.006)	(0.017)	(0.010)	(0.033)	(0.022)
Diff. (β_2-β_1)	0.07	0.00	-0.03	-0.01	-0.07*	-0.01
	(0.09)	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)
Model description:						
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.39	0.39	0.16	0.16	0.20	0.20
Control group mean	0.48	0.03	0.11	0.06	0.19	0.10
Observations	7736	7736	7519	7519	2767	2767

$$* : p < 0.10; \quad ** : p < 0.05; \quad *** : p < 0.01$$

Notes. Columns (1), (3), and (5) corresponds to the non-response rate in percent among sensitive questions (i.e. 0.1 means the respondent refused to answer 0.1 percent of sensitive questions).

Alternative version

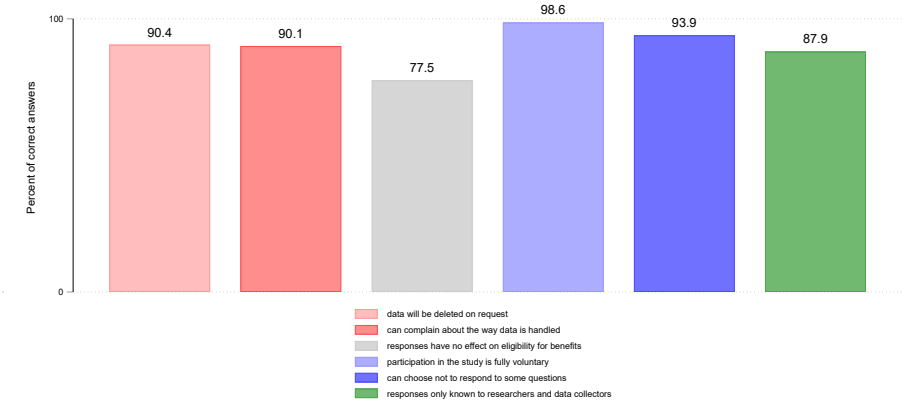
Conclusion

- We measured understanding related to informed consent and experimentally tested alternative approaches to inform respondents
- Overall respondents seem insufficiently informed by the business-as-usual consent process on key aspects
- Augmenting the process by showing a video decreased understanding of some aspects
- Adding an interactive reading of the consent form improved some aspects of understanding, but self-reported understanding is negatively impacted.
- There are no (adverse) effects on consent and no meaningful effects on item-non-response rates from the augmented approaches

$\Rightarrow$ Study participant's understanding of consent is insufficient but can be improved. Further research to develop better protocols is needed.

Appendix

Finding Ia-2: Limited understanding



Note: Share of correct responses when true is correct in benchmark group.

Figure: Objective measure: True correct

False correct version

Finding Ia-3a: Limited understanding

	Mean (1)	Obs. (2)
Knows purpose is only research	0.762	445
Purpose: needs assessment (determining the eligibility to a program)	0.182	445
Purpose: research	0.935	445
Purpose: marketing	0.027	445
Purpose: criminal prosecution	0.002	445
Knows participation is voluntary	0.760	445
If consent: I will be interviewed	0.847	445
If consent: my responses will be send to the researchers	0.227	445
If consent: my information will be saved for a pot. new int.	0.169	445
If consent: I will receive money or an in-kind compensation for the int.	0.002	445
If consent: a non-governmental organization will help me, my family, or my comm.	0.009	445
If no consent: the interview stops immediately	0.843	445
If no consent: I will be declined services in the future	0.103	445
If no consent: I will lose existing benefits	0.049	445
If no consent: someone will punish me	0.022	445

False correct version

Finding Ia-3b: Limited understanding

	Mean (1)	Obs. (2)
Knows data users are only researchers	0.764	445
Usage: researchers	0.901	445
Usage: non-governmental organization	0.144	445
Usage: private company	0.052	445
Usage: government	0.045	445
Knows data will not be shared with third parties	0.919	445
Has access: enumerators	0.578	445
Has access: researchers	0.789	445
Has access: data collection company	0.292	445
Has access: other private companies	0.007	445
Has access: non-governmental organizations	0.022	445
Has access: governmental or public institutions	0.022	445

False correct version

Finding Ia-4: Limited understanding

Table: Understanding in an Enumerator Pilot Study

	Overall (1)	Rights (2)	Purpose (3)	Voluntariness (4)	Confidentiality (5)
Objective understanding					
Share informed	2%	36% and 72%	29.6%	13% and 20%	99%
Subjective understanding					
Share understanding	73%	87%	94%	85%	93%
Objective and subjective					
High and high	1%	30%	29%	7%	92%
Low and high	72%	59%	65%	78%	1%
High and low	0%	4%	1%	0%	7%
Low and low	27%	7%	5%	15%	0%

False correct version

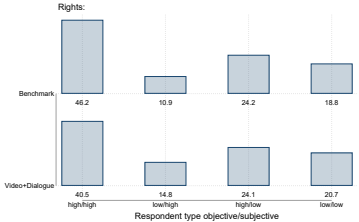
Findings III-2: Objective measure of understanding

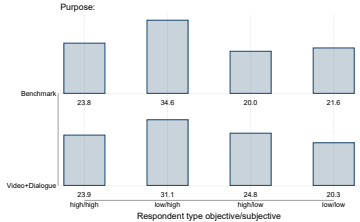
	Overall		Rights		Purpose	Voluntariness		Confidentiality
	(1) Score	(2) Informed	(3) Delete info.	(4) Can compl.	(5)	(6) Part. obliged	(7) Resp. obliged	(8)
Female	0.001	-0.044	0.046	-0.026	0.039	0.048	0.036	-0.000
	(0.019)	(0.041)	(0.044)	(0.040)	(0.032)	(0.036)	(0.038)	(0.039)
Video (β_1)			0.026	-0.109**	0.013	-0.040	0.026	-0.152***
			(0.054)	(0.052)	(0.033)	(0.036)	(0.036)	(0.056)
+Dialogue (β_2)	0.016	0.010	0.055	-0.067	0.032	0.038	0.110***	-0.044
	(0.018)	(0.039)	(0.069)	(0.058)	(0.040)	(0.037)	(0.039)	(0.062)
Video x Female (γ_1)			-0.061	-0.003	-0.067	0.025	-0.046	0.134**
			(0.064)	(0.066)	(0.042)	(0.047)	(0.049)	(0.066)
+Dialogue x Female (γ_2)	-0.011	0.034	-0.078	0.053	-0.053	0.026	-0.011	0.035
	(0.023)	(0.051)	(0.080)	(0.071)	(0.054)	(0.054)	(0.061)	(0.074)
Diff. (β_2-β_1)			0.03	0.04	0.02	0.08*	0.08**	0.11
			(0.07)	(0.07)	(0.04)	(0.04)	(0.04)	(0.07)
Diff. Female (γ_2-γ_1)			-0.02	0.06	0.01	0.00	0.03	-0.10
			(0.09)	(0.08)	(0.06)	(0.06)	(0.07)	(0.08)
Model description:								
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.58	0.53	0.48	0.53	0.45	0.51	0.46	0.51
Control group mean	0.67	0.38	0.59	0.51	0.19	0.29	0.34	0.69
Observations	888	888	807	819	1216	1225	1206	809

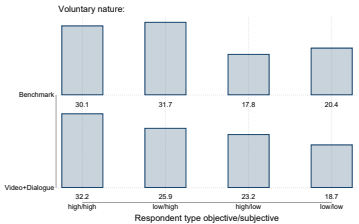
$* : p < 0.10;\ \ ** : p < 0.05;\ \ *** : p < 0.01$

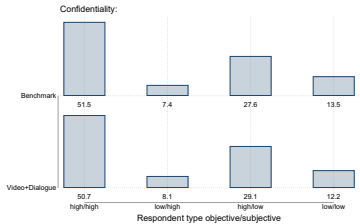
Main specification

Finding IV-2: Subjective vs. objective measure









Subjective measure

Findings V-2: Item non-response rates

	1st Visit		2nd Visit (HH Roster)		2nd Visit (Full Int.)	
	(1) Non-resp. rate	(2) Any non-resp.	(3) Non-resp. rate	(4) Any non-resp.	(5) Non-resp. rate	(6) Any non-resp.
Video (β_1)	-0.016	-0.002	-0.036	-0.018*	-0.069	-0.010
	(0.038)	(0.004)	(0.025)	(0.009)	(0.066)	(0.018)
+Dialogue (β_2)	-0.000	0.000	0.069	0.022	0.062	0.018
	(0.074)	(0.007)	(0.057)	(0.020)	(0.145)	(0.036)
Diff. (β_2-β_1)	0.02	0.00	0.10*	0.04**	0.13	0.03
	(0.07)	(0.01)	(0.06)	(0.02)	(0.15)	(0.04)
Model description:						
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.32	0.31	0.21	0.22	0.20	0.21
Control group mean	0.42	0.04	0.59	0.26	1.03	0.38
Observations	7736	7736	7519	7519	2767	2767

$$* : p < 0.10; \quad ** : p < 0.05; \quad *** : p < 0.01$$

Notes. Columns (1), (3), and (5) corresponds to the non-response rate in percent among sensitive questions (i.e. 0.1 means the respondent refused to answer 0.1 percent of sensitive questions).

Main version