Timmons et al. (2020). *It depends on how you ask: Measuring bias in population surveys of compliance with COVID-19 public health guidance.*

Supplementary Material

This supplement contains additional details of the participants and the design and analysis of the experiments.

Method

Participants

Table S1. *Sample Characteristics*

		Experiment 1 (Positive Frame; $N =$ 800)		Experiment 2 (Negative Frame; $N =$ 1000)		CSO Census of Population 2016
		n	%	*n*	%	%
Gender	Men	421	52.6	508	51	49.6
	Women	376	47	488	49	50.4
	Prefer Not To Say/Other	3	0.4	4	0.4	
Age	Under 40 years	250	31	360	36	38.3
	40 – 59 years	329	41	333	33	36.3
	60 years +	221	28	307	31	25.4
Education	Degree or above	315	39	436	44	42
	Below degree	485	61	564	56	58
Employment	Employed	455	57	563	56	51.9
	Not employed	345	43	437	44	48.1
Urban-Rural	Rural	296	37	373	37	39.1
	Urban	504	63	627	63	60.8

Materials, Design and Procedure

Participants were informed that in this section of the study they would be asked some general background questions, and about how the pandemic might have affected them and how they might be coping with it. They were reminded that all of their responses would be completely anonymous and it would not be possible to identify them.

They were then told that this stage of the study was about "some actions that people in Ireland might be taking during the pandemic" and presented with a list of items. They were asked to "Please note **how many** of the actions apply to you – you don't need to select which ones, just **how many**." Participants gave their responses using numbered buttons.

List items are presented in Table S2. In line with best practice for list experiments, non-target items were designed such that within each set of three, one would apply to a majority of participants (e.g. using the phone or internet to stay in touch with others), one would apply to very few participants (e.g. watching TV less often, as indicated by other surveys at the time), and one would apply to a mix of participants (e.g. spending more time with a household pet). Alternatively, two of the items were designed to negatively correlate (e.g. "I have had all of my groceries delivered to my home" and "I have noticed my local supermarket provides hand sanitiser"), with one item that would apply to a mix of participants. This design limits the likelihood that participants respond with minimum or maximum answers to the list questions, thereby encouraging anonymity in endorsement of the target item. (See Reference [6] in the main text.)

Table S2. *Target and Non-Target List Items*

Set 1	Target	I wash my hands with soap and water for a full 20 seconds (or I use hand sanitiser) when I return home from being out or touch a surface other people might have touched
	Non-Target	I have been keeping in touch with friends and family via the internet or by phone
		I am watching less TV (or streaming TV shows less) than usual
		I have a household pet that I have been spending more time with (e.g. taking the dog for a walk more often)
Set 2	Target	I keep at least 2m between myself and people outside of my household
	Non-Target	I have had all of my groceries delivered to my home
		I have noticed my local supermarket provides hand sanitiser at or near the door
		I have purchased food from a local restaurant for collection or delivery about once per week
Set 3	Target	*Experiment 1:* I had limited face-to-face contact with people outside my household before June 8th (e.g. never more than 3 other people and always outdoors)
		Experiment 2: I **don't** usually wear a face-covering or mask when in public places (e.g. the supermarket, shops and public transport)
	Non-Target	I have watched, listened to or read news updates on the COVID-19 pandemic, at least a couple of times per week
		I have been spending more money than usual during the pandemic
		I have been exercising more often than normal

Note. The frame of all items were reversed for Experiment 2.

Results

Data were analysed using the *kict* package in Stata[1]. Prior to each analysis, responses were checked for the assumption of no design effect; no set of items showed significant effects, suggesting no evidence of design effects. We use least squares estimators, although closely similar results are found using maximum likelihood estimators.

Meeting Others and Mask-Wearing

In Experiment 1, more participants reported only limited contact with others when asked directly (83.4%) than in the list (74.2%), $t(494.0) = 1.59$, $p = .056$, $d = .11$. In Experiment 2, more participants reported usually wearing a mask in public in the List condition (67.6%) than the Direct condition (54.2%), $t(659.9) = 2.24$, $p = .013$, $d = 0.14$. This effect is the converse of our prediction, implying that participants judged *not* wearing a mask to be socially desirable. However, one possibility is that adding the adverb 'usually' to this item (noted above) interacted with the list experiment design and altered interpretation. Significant differences by socio-demographic subgroups are reported in Table S3.

Table S3.

Significant Differences by Socio-Demographic Subgroups

Health Behaviour	Socio-Demographic Characteristic	Finding
Meeting Others	Age	Younger participants (aged under 50 years) exhibited social desirability bias for meeting others (81.8% vs. 69.9%), $t(246.8) = 1.43$, $p = .077$, $d = 0.14$.
	Living Area	Rural participants displayed significant social desirability bias when asked about meeting others (84.2% vs. 66.8%), $t(187.8) = 1.95$, $p = .026$, $d = 0.23$.

Mask-wearing	Gender	Fewer men reported wearing a mask when asked Directly than when asked in the List condition (49.1% vs. 79.1%), t (317.87) = 2.45, p = .007, d = 0.22.

Response Time Exclusions

The same analyses as those reported in the main text were repeated with the fastest 5% of participants trimmed on each set, with results presented in Table S4. The pattern of results is the same as those reported in the main text.

Table S4. *Analyses with Response Time Exclusions*

Item	Experiment	Condition	%	n	t	df	p
Hand Washing	Experiment 1	Direct	91.3%	379	1.42	432.4	.078
		List	83.1%	382			
	Experiment 2	Direct	83.6%	472	0.03	572.6	.488
		List	83.5%	470			
Distancing	Experiment 1	Direct	92.3%	379	0.69	443.1	.246
		List	88.9%	382			
	Experiment 2	Direct	79.5%	473	1.02	571.1	.153
		List	73.2%	477			
Meeting Others	Experiment 1	Direct	83.4%	379	1.65	472.6	.050
		List	73.8%	383			
Mask Wearing	Experiment 2	Direct	55.1%	474	2.19	628.1	.014
		List	68.5%	478			

Further Exclusions: Maximum and Minimum Responding

Table S5 shows the percentage of participants in the List condition who responded with a '0' or '4' on each set of items (meaning their endorsement or lack thereof of the target item would have been non-anonymous). Proportions were somewhat higher when items were framed negatively, particularly among those who gave a '0' response.

Table S5. *Percentage of Maximum and Minimum Responses in Both Experiments*

	Experiment 1		Experiment 2	
	Minimum	Maximum	Minimum	Maximum
Hand-washing	1.0%	7.0%	22.5%	1%
Distancing	1.2%	4.7%	15.1%	1.4%

Running the same analysis on hand-washing with exclusions for the large minority who responded with a '0' showed a significant effect of social desirability (83.1% vs. 78.2%), $t(751.28) = 1.34$, $p = .091$, $d = 0.09$, replicating Experiment 1. Distancing was not significant (79.3% vs. 76.6%), $t(571.09) = 1.02$, $p = .153$, $d = 0.06$, again replicating Experiment 1.

---

[1] Tsai CL. Statistical analysis of the item-count technique using Stata. The Stata Journal. 2019 Jun;19(2):390-434.