

Social norms and covid vaccines: Analysis plan

The Policy Lab at Brown University

randomize{Nat Rabb, Jake Bowers, David Glick, Attiya Houston, David Yokum}
February 19, 2021

We reason from past literature to hypothesize that beliefs about the vaccine intentions of others predict people's own vaccine intentions. If we are right, and if beliefs about the intentions of others are malleable, then we will have found an effective route for messaging about vaccinations.

Prominent **alternative explanations** for vaccine intention include: political ideology and party ID (which reflects beliefs about descriptive norms about vaccinations too perhaps), factors which indicate objective risk of more severe Covid-19 outcomes (such as age, gender, race, and income), and factors which tend to predict differences in health behaviors in general (such as education).

Although it would be difficult to change the elite messaging around vaccination that people are choosing to absorb, we could in principle **tell people what others have said** using survey data (sharing it openly) and/or vivid interviews. Thus, there is a policy reason to pursue the idea of social norms and social proof.

Using survey data collected at four time points between August and December, 2020 (<https://osf.io/w2nhg>), we will answer the following research questions with the corresponding analysis methods.

1 Primary Hypotheses and Analyses

Research question	Analysis	Populations	Notes
Does perceived others' intentions to vaccinate predict vaccination intention?	Multiple regression after Matching Outcome: vaccine intention Focal Var: <u>others' intentions:</u> family/friends' vaccine intention, neighbors' vaccine intention, people in city vaccine intention, people in state vaccine intention Controls: Alternative	Survey 3 RI ($N = 500$) Survey 3 US ($N = 1000$) Survey 4 RI ($N = 500$)	How many others' intention variables do we have? Should we combine them? Or look at them each individually?

	explanations listed above		
Does perceived others' intentions to get tested predict willingness to get tested better?	<p>Multiple regression after Matching</p> <p>Outcome: willingness to get tested</p> <p>Focal Var: family/friends' intention to get tested</p> <p>Controls: Alternative explanations listed above</p>	<p>Survey 1 RI ($N = 500$)</p> <p>Survey 1 US ($N = 1000$)</p> <p>Survey 2 RI ($N = 696$)</p> <p>Survey 2 oversample ($N = 1450$)</p> <p>Survey 3 RI ($N = 500$)</p> <p>Survey 3 oversample ($N = 295$)</p> <p>Survey 3 US ($N = 1000$)</p> <p>Survey 4 RI ($N = 500$)</p>	The response scale changed from Surveys 1 & 2 to Surveys 3 & 4 to fix a problem with interpretation (the first two surveys showed numeric ranges rather than proportions, thus failing to account for varying family sizes)
Does perceived others' intentions to wear masks predict their own mask wearing intentions?	<p>Multiple regression after Matching</p> <p>Outcome: mask-wearing when visiting family & friends</p> <p>Focal var: family/friends' mask-wearing</p> <p>Controls: Alternative explanations listed above</p>	Survey 4 RI ($N = 500$)	

1.1 Details of the research design:

We will create strata of survey respondents who are similar in terms of the variables that we label “control” variables above but who differ as much as possible in terms of their perceptions of the intentions of others. We will not look at outcomes until we are satisfied with our strata — and plan to start by using optimal full matching (in the optmatch R package) but may also use optimal pair matching with constraints on covariate differences (in the designmatch R package) if we find it difficult to create a research design that (1) keeps as many survey respondents as possible and (2) creates strata that, collectively, convince us that we are comparing like with like (using standards such as the block randomized experiment in the Rltools R package).

The control variables are the following (using names from MERGE_NR_2.3.21 dataset):

Dimension	Variable	Response options
Political ideology	ideo5	<1>"Very liberal"<2>"Liberal"<3>"Moderate"<4>"Conservative"<5>"Very conservative"
Party ID	pid3	<1>Democrat<2>Republican<3>Independent
Age	age	[continuous]
Gender	gender	<1>Male<2>Female<3>Other<4>Prefer

		not to say
Race	race	<1>White<2>Black<3>Hispanic<4>Asian<5>Native American<6>Mixed<7>Other<8>Middle Eastern
Income	faminc_new	<1>"Less than \$10,000"<2>"\$10,000 - \$19,999" <3>"\$20,000 - \$29,999" <4>"\$30,000 - \$39,999"<5>"\$40,000 - \$49,999" <6>"\$50,000 - \$59,999"<7>"\$60,000 - \$69,999" <8>"\$70,000 - \$79,999"<9>"\$80,000 - \$99,999" <10>"\$100,000 - \$119,999"<11>"\$120,000 - \$149,999"<12>"\$150,000 - \$199,999"<13>"\$200,000 - \$249,999"<14>"\$250,000 - \$349,999"<15>"\$350,000 - \$499,999"<16>"\$500,000 or more"
Education	educ	<1>"No HS"<2>"High school graduate"<3>"Some college"<4>"2-year"<5>"4-year"<6>"Post-grad"

And the key outcome and focal variable names are:

Dimension	Variable	Response options
Intention to get vaccinated	q1	<1>Definitely won't<2>Maybe won't<3>Not sure<4>Maybe will<5>Definitely will
Perceived friends'/family members' intentions to get vaccinated	q60_1	<1>None<2>A few (less than a quarter)<3>About a quarter<4>About half<5>About three quarters<6>Most (more than three quarters)<7>All
Perceived neighbors' intentions to get vaccinated	q60_2	"
Perceived people in your city/town's intention to get vaccinated	q60_3	"
Perceived people in your state's intention to get vaccinated	q60_4	"
Willingness to get tested	q29	<1>Absolutely not<2>Probably unwilling<3>Somewhat unwilling<4>Somewhat willing<5>Probably willing<6>Absolutely
Perceived friends'/family members' willingness to get tested	q25_number* q25	q25_number: <1>0<2>1-4<3>5-9<4>10-14<5>15 or more

		<p>*This scale neglects the denominator (size of one's social network) so for half of the surveys we switched to q25:</p> <p><1>None<2>A few (less than a quarter)<3>About a quarter<4>About half<5>About three quarters<6>Most (more than three quarters)<7>All</p>
Mask-wearing around friends & family	q7	<p><1>Never<2>Not very often<3>Sometimes<4>Very often<5>Every time</p>
Perceived friends'/family members' mask-wearing around you	q100	"

1.2 Details of the statistical analysis:

We are interested in (1) whether we can detect any difference between groups who perceive more or less readiness to vaccinate, test, or wear masks in terms of their own stated intentions (so we will deploy hypothesis tests, probably focusing on both means as well as ranks in regards test statistics. We may try using *both* means and ranks as a way to increase statistical power following Rosenbaum 2012 (the functions for adaptive testing in Rosenbaum's r packages) and (2) the magnitude of the difference between groups (so we will estimate this difference using differences of means calculated using linear least squares but using HC2 standard errors following the design-based justification of Winston Lin, Joel Middleton, Peter Aronow, etc.)

We do not plan to adjust p-values for multiple comparisons: a "significant" finding for one outcome stands on its own, we do not see the different outcomes as a single family of tests.

2 Planned Descriptive/Exploratory/Follow-Up Analyses

We plan to execute the following analyses in the process of writing the paper.

1. If people do not vary in their perceptions of others' intentions, then we have no variance in our focal variable that can be used to address our scientific question. So, we describe this variable.
2. Rarely has the scientific community seen simple descriptions of perceptions of others' intentions for vaccination and other health behaviors: This is another reason we present a table of descriptive information early in the paper.
3. We also plan to look at **accuracy** of these perceptions and to ask whether it is **perceived level** or **accuracy** which matters most from the perspective of predicting intentions about their own behavior. This is exploratory because we know of little past literature that we can build on to develop hypotheses about accuracy and behavior.

However, these analyses will be useful in regards to setting up future studies and adding to the scientific and policy learning about these topics.

How accurate are people's beliefs about others' vaccination intentions?	Descriptives of % statewide that will definitely get vaccine and % that participants estimated will definitely get the vaccine	Survey 3 RI ($N = 500$) Survey 3 oversample ($N = 295$) Survey 3 US ($N = 1000$) Survey 4 RI ($N = 500$)	
Does accuracy of beliefs vary by demographics?	Estimates by demographics	Survey 3 RI ($N = 500$) Survey 3 oversample ($N = 295$) Survey 3 US ($N = 1000$) Survey 4 RI ($N = 500$)	
Is perceived statewide intention to vaccinate or accuracy of perceived statewide intention to vaccinate a better predictor of own intention to vaccinate?	Regression DV: vaccine intention IVs: people in state vaccine intention, accuracy of same	Survey 3 RI ($N = 500$) Survey 3 oversample ($N = 295$) Survey 3 US ($N = 1000$) * Survey 4 RI ($N = 500$)	
Does the magnitude of the gap between perceived friends/family and perceived statewide intentions to vaccinate predict one's own vaccine intention?	Correlation	Survey 3 RI ($N = 500$) Survey 3 oversample ($N = 295$) Survey 3 US ($N = 1000$) * Survey 4 RI ($N = 500$)	
Does the magnitude of the gap between perceived friends/family and perceived statewide intentions to vaccinate vary by demographics?	Descriptives of difference scores by demographics	Survey 3 RI ($N = 500$) Survey 3 oversample ($N = 295$) Survey 3 US ($N = 1000$) Survey 4 RI ($N = 500$)	
Did question order (others' intentions before or after own intention) affect vaccine intention?	Cochran-Mantel-Haenszel test (b/c blocked on party ID) for main effect Parallel test for effects on people with extreme family/friends' ratings (1 or 7)	Survey 4 RI ($N = 500$)	

The additional variables for these analyses are:

Dimension	Operationalization
Accuracy (by state)	Accurate = Response chosen on q60_4 is nearest neighbor to true value (e.g., if the proportion of people in a state that will definitely get vaccinated is 37%, both "About a quarter" and "About half" would be considered accurate). True values will be determined by multilevel

	regression and poststratification or independent estimates (e.g., from Understanding Coronavirus in America)
Accuracy (by city/town)	Accurate = Response chosen on q60_3 is nearest neighbor to true value (e.g., if the proportion of people in a state that will definitely get vaccinated is 37%, both “About a quarter” and “About half” would be considered accurate). True values will be determined by multilevel regression and poststratification.
Magnitude of gap between perceived intentions to vaccinate of innermost and outermost circles of others	Difference score: q60_4 – q60_1
Question order	q60_treat (1 = people rated their perceptions of others’ intentions to vaccinate before rating their own; 2 = people rated others’ intentions after their own)

3 Other Planned Decisions: Missing data, Extreme data.

We do not anticipate any extreme data.

We will replace covariate or “control” missing data with means and add an indicator for missingness (0=not missing, 1=missing) for each such variable with missing data. We will match on *both* the values of the covariates *and* the missingness of the covariates.

We will drop any observations with missing outcomes or missing values on the focal variables. We will report a table that compares these observations with the observations that we use in the analysis in the Appendix so that readers can judge the generalizability of our findings: in the extreme case, imagine that we have to drop all men from the analysis, this table will alert us that our inferences may not teach us much about the effects of perceived others’ intentions on men’s intentions.