



A 680,000-person megastudy of nudges to encourage vaccination in pharmacies

Katherine L. Milkman^{a,1}, Linnea Gandhi^a, Mitesh S. Patel^b, Heather N. Graci^c, Dena M. Gromet^c, Hung Ho^d, Joseph S. Kay^e, Timothy W. Lee^e, Jake Rothschild^c, Jonathan E. Bogard^f, Ilana Brody^f, Christopher F. Chabris^g, Edward Chang^h, Gretchen B. Chapmanⁱ, Jennifer E. Dannals^j, Noah J. Goldstein^k, Amir Goren^g, Hal Hershfield^l, Alex Hirsch^m, Jillian Hmurovicⁿ, Samantha Hornⁱ, Dean S. Karlan^o, Ariella S. Kristal^p, Cait Lambertson^q, Michelle N. Meyer^g, Allison H. Oakes^r, Maurice E. Schweitzer^a, Maheen Shermohammed^g, Joachim Talloenⁱ, Caleb Warren^s, Ashley Whillans^h, Kuldeep N. Yadav^t, Julian J. Zlatev^h, Ron Berhan^q, Chalanda N. Evans^u, Rahul Ladhania^v, Jens Ludwig^w, Nina Mazar^x, Sendhil Mullainathan^y, Christopher K. Snider^z, Jann Spiess^{aa}, Eli Tsukayama^{bb}, Lyle Ungar^m, Christophe Van den Bulte^q, Kevin G. Volpp^{cc}, and Angela L. Duckworth^{ad}

^aDepartment of Operations, Information and Decisions, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; ^bAscension Health, St. Louis, MO 63105; ^cBehavior Change for Good Initiative, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; ^dDepartment of Marketing, The University of Chicago Booth School of Business, Chicago, IL 60637; ^eSchool of Professional Studies, Northwestern University, Evanston, IL 60208; ^fDepartment of Behavioral Decision Making, Anderson School of Management, University of California, Los Angeles, CA 90095; ^gBehavioral Insights Team, Steele Institute for Health Innovation, Geisinger Health System, Danville, PA 17822; ^hNegotiation, Organizations & Markets Unit, Harvard Business School, Harvard University, Boston, MA 02163; ⁱDepartment of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213; ^jDepartment of Strategy and Management, Tuck Business School, Dartmouth College, Hanover, NH 15213; ^kDepartment of Management and Organizations, Anderson School of Management, University of California, Los Angeles, CA 90095; ^lDepartment of Marketing, Anderson School of Management, University of California, Los Angeles, CA 90095; ^mDepartment of Computer and Information Science, School of Engineering, University of Pennsylvania, Philadelphia, PA 19104; ⁿWharton Risk Management and Decision Processes Center, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; ^oDepartment of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; ^pDepartment of Organizational Behavior, Harvard Business School, Boston, MA 02163; ^qDepartment of Marketing, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; ^rAnthem, Inc., Indianapolis, IN 46203; ^sDepartment of Marketing, Eller College of Management, University of Arizona, Tucson, AZ 85721; ^tDepartment of Biostatistics, Epidemiology, & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; ^uPenn Medicine Nudge Unit, Penn Medicine, Philadelphia, PA 19104; ^vDepartment of Health Management and Policy, University of Michigan School of Public Health, Ann Arbor, MI 48109; ^wHarris School of Public Policy, University of Chicago, Chicago, IL 60627; ^xDepartment of Marketing, Questrom School of Business, Boston University, Boston, MA 02215; ^yDepartment of Behavioral Science, Booth School of Business, University of Chicago, Chicago, IL 60637; ^zCenter for Health Care Innovation, Penn Medicine, Philadelphia, PA 19104; ^{aa}Department of Operations, Information, and Technology, Graduate School of Business, Stanford University, Stanford, CA 94305; ^{bb}Business Administration Division, University of Hawaii–West Oahu, Kapolei, HI 96707; ^{cc}Penn Center for Health Incentives and Behavioral Economics, Departments of Medical Ethics and Health Policy and Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; and ^{ad}Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104

Edited by Elke Weber, Departments of Psychology and Public Affairs, Princeton University, Princeton, NJ; received August 16, 2021; accepted December 20, 2021

Encouraging vaccination is a pressing policy problem. To assess whether text-based reminders can encourage pharmacy vaccination and what kinds of messages work best, we conducted a megastudy. We randomly assigned 689,693 Walmart pharmacy patients to receive one of 22 different text reminders using a variety of different behavioral science principles to nudge flu vaccination or to a business-as-usual control condition that received no messages. We found that the reminder texts that we tested increased pharmacy vaccination rates by an average of 2.0 percentage points, or 6.8%, over a 3-mo follow-up period. The most-effective messages reminded patients that a flu shot was waiting for them and delivered reminders on multiple days. The top-performing intervention included two texts delivered 3 d apart and communicated to patients that a vaccine was “waiting for you.” Neither experts nor lay people anticipated that this would be the best-performing treatment, underscoring the value of simultaneously testing many different nudges in a highly powered megastudy.

Significance

Encouraging vaccination is a pressing policy problem. Our megastudy with 689,693 Walmart pharmacy customers demonstrates that text-based reminders can encourage pharmacy vaccination and establishes what kinds of messages work best. We tested 22 different text reminders using a variety of different behavioral science principles to nudge flu vaccination. Reminder texts increased vaccination rates by an average of 2.0 percentage points (6.8%) over a business-as-usual control condition. The most-effective messages reminded patients that a flu shot was waiting for them and delivered reminders on multiple days. The top-performing intervention included two texts 3 d apart and stated that a vaccine was “waiting for you.” Forecasters failed to anticipate that this would be the best-performing treatment, underscoring the value of testing.

vaccination | COVID-19 | nudge | influenza | field experiment

Encouraging vaccination has emerged as a pressing policy problem during the COVID-19 crisis. Although COVID-19 vaccines were widely available to Americans beginning in the late spring of 2021, the United States did not achieve President Biden’s goal of 70% of American adults receiving their first dose by July 1 (1). Further, millions who received a first dose of a COVID-19 vaccine failed to receive a second dose on schedule (1). States, cities, and the federal government have been left scrambling to find cost-effective ways to motivate vaccine uptake (2, 3).

How can policymakers more-effectively encourage vaccination? Past research suggests that simple, low-cost nudges can help (4–8). For instance, when doctors’ offices and health systems send text-based reminders to patients that vaccines are

Author contributions: K.L.M., L.G., M.S.P., D.M.G., J.E.B., I.B., C.F.C., E.C., G.B.C., J.E.D., N.J.G., A.G., H. Hershfield, A.H., J.H., S.H., D.S.K., A.S.K., C.L., M.N.M., A.H.O., M.E.S., M.S., J.T., C.W., A.W., K.N.Y., J.J.Z., R.L., J.L., N.M., S.M., J.S., L.U., and A.L.D. designed research; K.L.M., L.G., M.S.P., H.N.G., D.M.G., H. Ho, J.S.K., T.W.L., and J.R. performed research; R.L., J.L., N.M., S.M., J.S., E.T., and L.U. contributed new reagents/analytic tools; L.G., H.N.G., H. Ho, T.W.L., J.R., C.N.E., and C.K.S. analyzed data; K.L.M., L.G., M.S.P., J.S.K., R.B., E.T., C.V.d.B., and K.G.V. supervised data analysis; K.L.M., L.G., M.S.P., H.N.G., D.M.G., H. Ho, T.W.L., J.R., and J.S.K. prepared the supplementary information; K.L.M., L.G., M.S.P., D.M.G., R.B., C.V.d.B., K.G.V., and A.L.D. wrote the paper; and M.S.P., H.N.G., H. Ho, T.W.L., J.R., J.S.K., R.B., E.T., C.V.d.B., and K.G.V. provided critical feedback on the paper.

Competing interest statement: K.G.V. is a part-owner of VAL Health, a behavioral economics consulting firm.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: kmilkman@wharton.upenn.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2115126119/-DCSupplemental>.

Published February 1, 2022.

available or reserved for them, this significantly boosts both flu vaccination (8) and COVID-19 vaccination rates (9).

Increasing compliance with public health recommendations is not just a persuasion problem. Decades of research show that many people fail to follow through on their intentions when it comes to decisions about their health (10). In the context of vaccination, one study found that only 79% of those who intended to get a flu shot actually followed through (11). Follow-through failures are driven by a combination of forgetting, failing to anticipate and plan for obstacles, and low motivation (10, 12). Psychologically informed reminders have the potential to bridge some of these “intention–action” gaps (13), and pharmacies have become a front line for providing vaccinations (14), providing 37% of flu shots in the United States as of 2020 (15). As a result, understanding what reminders work in this context is particularly important, especially because patient demographics differ in pharmacies and doctors’ offices (16).

Although there are many important differences between COVID-19 and influenza, both are respiratory ailments for which the risk of infection and severe illness can be reduced with safe, widely available vaccines. In the fall of 2020, anticipating the pressing need for insights about behaviorally informed COVID-19 vaccination messaging, we partnered with Walmart pharmacies to conduct a megastudy (17) of text-based reminder messages to encourage flu vaccination.

Pharmacy patients ($n = 689,693$) were randomly assigned to receive one of 22 different text message reminders or to a business-as-usual control condition with no messages. Text messages were developed by separate teams of behavioral scientists and used a variety of different tactics to nudge vaccination.

While all 22 interventions were designed to encourage inoculation against the flu, scientists were asked to develop interventions that they believed would be reusable to encourage COVID-19 vaccination. The resulting interventions relied on a wide range of behavioral insights and varied both in their content and timing of delivery.

To assess how well the relative success of these messages could be forecasted ex ante, both the scientists who developed the texts and a separate sample of lay survey respondents predicted the impact of different interventions on flu vaccination rates.

Megastudy Methods

We conducted our megastudy in partnership with Walmart, a large retail corporation with over 4,700 pharmacies across the United States. The study took place in the fall and winter of 2020. Walmart pharmacy patients in the United States were eligible for inclusion if they had not received a flu shot at Walmart by September 25, 2020, had received a flu shot at a Walmart during the 2019 to 2020 flu season, had consented to receive short messaging service (SMS; commonly known as text messaging) communications from Walmart pharmacy, and did not share a phone number with another eligible patient. This yielded a sample of 689,693 patients (23.7% of the 2,906,701 total patients who received a flu shot at a Walmart pharmacy between June 1, 2019, and May 31, 2020).

Patients were randomly assigned to one of our 22 different treatment conditions or a business-as-usual control condition in which patients didn’t receive any reminder messages to get vaccinated, following Walmart’s business-as-usual protocol. Randomization was conducted with the Java random number generator in the Apache Hadoop software library (<https://hadoop.apache.org/>).^{*} At least 27,365 patients were assigned to each study condition (average $n = 29,987$; median $n = 27,715$).

All 22 text-messaging interventions were designed by independent scientific teams to test distinct hypotheses about how

to encourage patients to get a flu shot at Walmart. For example, one study used humor in an attempt to increase memorability (“Did you hear the joke about the flu? Never mind, we don’t want to spread it around”) (18). Another communicated that vaccination was a growing social norm (i.e., “More Americans are getting the flu shot than ever”) (19, 20). Yet another intervention sought to increase goal commitment by prompting patients to text back if they planned to get a vaccine (“If you plan to get a flu shot at Walmart, commit by texting back: I will get a flu shot”) (21).

Intervention content differed on several attributes. Some were interactive, requesting that the patient text back a reply, and others were not. Some were more verbose; others were brief. To classify message content, we first conducted preregistered attribute analyses to identify the underlying characteristics of text messages that could help explain their efficacy (*SI Appendix*). We recruited 2,081 people from Prolific to rate text message content on five subjective dimensions (e.g., surprising). We also coded text messages on 10 objective dimensions (e.g., word count). We then analyzed bivariate correlations between each of these 15 attributes and intervention efficacy. To account for the nonindependence of attribute ratings, we performed a principal component analysis on the six attributes whose associations with efficacy were statistically significant ($P < 0.05$; *SI Appendix, Table S12*).

Notably, interventions differed not only in their content, but also in the frequency of message delivery: some texts were sent on one day only, while others were sent on two separate days. All intervention messages are provided in *SI Appendix*. Walmart sent the first message in each intervention on September 25, 2020, at 6:10 PM local time. For the 13 interventions that involved a second text message, it was delivered up to 3 d later (i.e., September 28, 2020).

Our study design and analyses were preregistered on AsPredicted (1: <https://aspredicted.org/blind.php?x=g39h9p>, 2: <https://aspredicted.org/blind.php?x=e2qh3d>, 3: <https://aspredicted.org/blind.php?x=29wz2i>).[†] A detailed description of our protocol is included in *SI Appendix*. This research was approved by the Institutional Review Board (IRB) at the University of Pennsylvania. The IRB granted a waiver of consent for this research. No identifying information about study patients was shared with the researchers.

Megastudy Results

Walmart pharmacy patients in our study were an average of 60.4 y old (SD = 15.8), and 61.6% were female. Data on the ethnicity of 84.2% of the patients in our sample were unavailable, but of the 15.8% of patients with available data on ethnicity, 80.3% were White, 8.7% were Hispanic, 7.6% were Black, and 3.3% were Asian. Patients came from all 50 US states (*SI Appendix, Table S2*). As shown in *SI Appendix, Table S1*, our 22 study conditions were balanced on age, race, and gender (P values from all F-tests > 0.05).

As preregistered, our primary outcome measure was whether patients in our study received their flu shot at a Walmart pharmacy during the 3-mo period between September 25, 2020 (when interventions began), and December 31, 2020. Our secondary outcome measure was whether patients in our study received their flu shot at a Walmart pharmacy during the 1-mo period between September 25, 2020, and October 31, 2020. Both outcomes were based on Walmart pharmacy records.

Following our preregistered analysis plan, we used an ordinary least-squares (OLS) regression to predict vaccination. The

^{*}Due to a coding bug, patients were assigned randomly but with unequal probabilities to the different study conditions. Patients were assigned to eight experimental conditions with a 1/20 probability and to the other 14 experimental conditions and the business-as-usual control condition with a 1/25 probability.

[†]Note that preregistration 2 makes small updates to preregistration 1, both of which were posted before any data were analyzed. Preregistration 3 details our data collection and analysis plan for the message content attributes analysis.

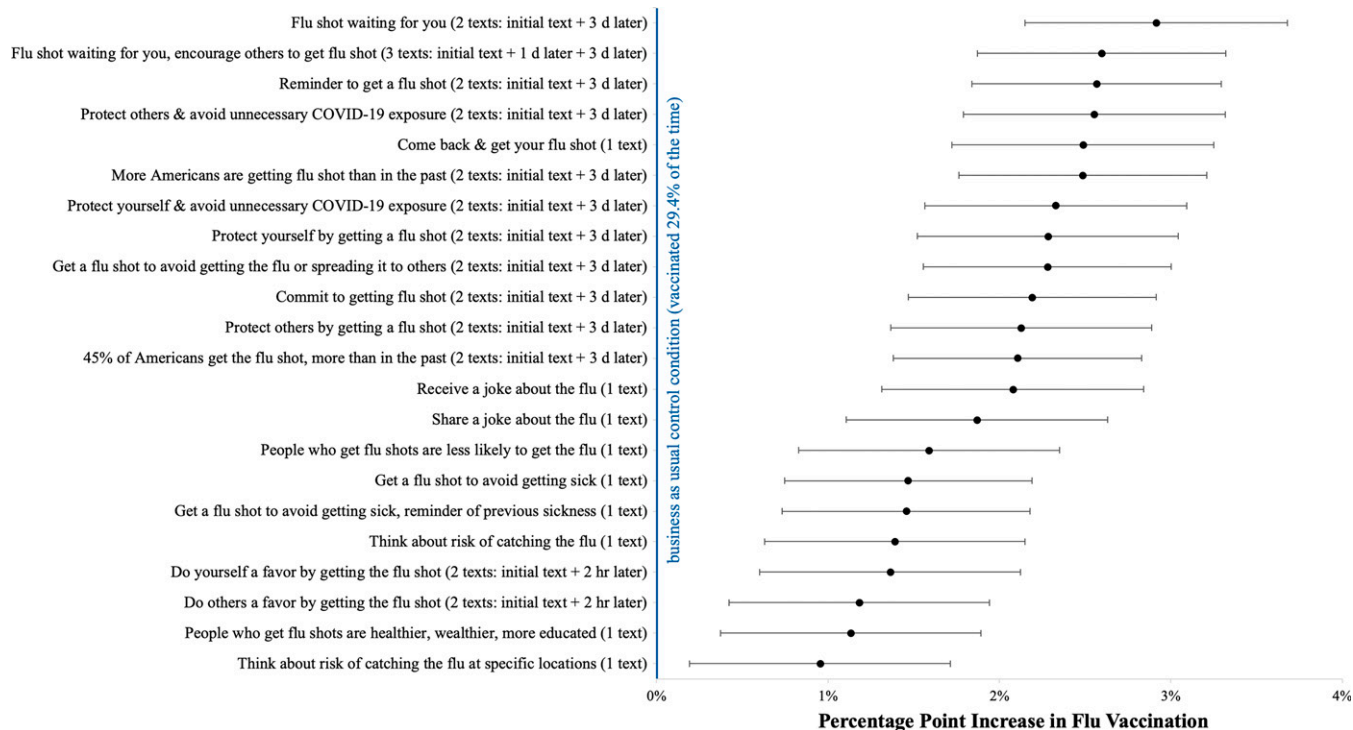


Fig. 1. Regression-estimated impact of each of our megastudy's 22 intervention conditions on flu vaccine uptake at Walmart by December 31st, 2020. Whiskers depict 95% CIs without correction for multiple comparisons.

predictors in our regression were 22 indicators for assignment to each of our study's experimental conditions with an indicator for assignment to our business-as-usual control condition omitted. Our preregistered regression included the following control variables: 1) patient age, 2) an indicator for whether a patient was male, and 3) indicators for patient race/ethnicity (Black non-Hispanic, Hispanic, Asian, and other/unknown; white non-Hispanic omitted). Because data on patient ethnicity was unexpectedly missing for 84.2% of our sample, we added controls to our regression for the fraction of citizens in a patient's home county who were white, Black, and Hispanic, according to the 2012 to 2016 American Community Survey (22). Our results are robust to excluding all controls (*SI Appendix, Table S10*).

In total, 29.4% of patients in the business-as-usual control condition received a flu shot at a Walmart pharmacy between September 25 and December 31, 2020. Fig. 1 and Table 1 report the results of our preregistered regression model predicting flu vaccination during this 3-mo period. Each of the 22 interventions that we tested significantly increased vaccination rates relative to the business-as-usual control condition (all two-sided, unadjusted P s < 0.02). Across our 22 experimental conditions, the mean treatment effect was a 2.0 percentage point increase in flu vaccinations relative to the business-as-usual control condition (a 6.8% lift, unadjusted $P = 0.003$).

In addition to reporting unadjusted robust SEs and CIs, Table 1 reports adjusted P values computed using the Benjamini–Hochberg procedure, which controls for the false discovery rate when conducting multiple comparisons (23). Of the 22 estimates, 21 have adjusted P values of less than 0.01.

However, not all treatments were equally effective. That is, we can reject the null hypothesis that all 22 effects have the same true value ($\chi^2 = 92.486$, $df = 21$, and $P < 0.001$).

As shown in Fig. 2, relative to the business-as-usual control condition, the top-performing intervention produced a 2.9 percentage point increase in flu vaccinations (a 9.9% lift; unadjusted and adjusted $P < 0.001$). After applying the James–Stein

shrinkage procedure to account for the fact that the maximum of 22 effects is likely to be upward biased (the winner's curse), we still estimate that this intervention produced a 2.7 percentage point increase in vaccination over the business-as-usual control condition (a 9.3% lift) (24). This intervention included two text messages. The first conveyed: "It's flu season & you can get a flu shot at Walmart." The second message arrived 72 h later and reminded patients that "A flu shot is waiting for you at Walmart."

Results for our secondary outcome variable—vaccination on or before October 31, 2020—are similar. As shown in *SI Appendix, Table S2*, all 22 experimental conditions significantly increased vaccination rates relative to the business-as-usual control condition, and they did so by an average of 2.1 percentage points (an 8.9% lift; all unadjusted P s < 0.02). Further, the same intervention (which described the vaccine as "waiting for you") was the top performer, increasing the regression-estimated vaccination rate among patients studied in the roughly 1 mo after contact by 3.2 percentage points (a 13.6% lift, unadjusted $P < 0.001$) or 3.0 percentage points (a 12.9% lift) after applying the James–Stein shrinkage procedure.

Principal component analysis revealed two underlying attributes of message content: incongruence with typical retail pharmacy communications (interrogative, surprising, imperative, and negative mood) and reminders that implied that the decision to get a shot had already been made and that a shot awaited them (waiting for you; reminder).

To examine how these two key content attributes and the timing of message delivery related to intervention efficacy, we then ran a preregistered OLS regression predicting each intervention's estimated treatment effect (i.e., the 22 coefficient estimates from Table 1). The predictors in this regression were our two content attributes as well as an indicator for whether intervention messages were sent on multiple days (as opposed to on 1 d), and we calculated heteroskedasticity-robust SEs. We found that the most-effective interventions 1) were framed as reminders to get flu shots that were already waiting for patients

Table 1. Regression-estimated impact of each of our megastudy's 22 intervention conditions on flu vaccine uptake at Walmart by December 31st, 2020

	Beta	SE	P value	Adjusted P value
Flu shot waiting for you (two texts: initial text + 3 d later)	0.029	(0.004)	<0.001	<0.001
Flu shot waiting for you, encourage others to get flu shot (three texts: initial text + 1 d later + 3 d later)	0.026	(0.004)	<0.001	<0.001
Reminder to get a flu shot (two texts: initial text + 3 d later)	0.026	(0.004)	<0.001	<0.001
Protect others and avoid unnecessary COVID-19 exposure (two texts: initial text + 3 d later)	0.026	(0.004)	<0.001	<0.001
More Americans are getting flu shot than in the past (two texts: initial text + 3 d later)	0.025	(0.004)	<0.001	<0.001
Come back and get your flu shot (one text)	0.025	(0.004)	<0.001	<0.001
Protect yourself and avoid unnecessary COVID-19 exposure (two texts: initial text + 3 d later)	0.023	(0.004)	<0.001	<0.001
Protect yourself by getting a flu shot (two texts: initial text + 3 d later)	0.023	(0.004)	<0.001	<0.001
Get a flu shot to avoid getting the flu or spreading it to others (two texts: initial text + 3 d later)	0.023	(0.004)	<0.001	<0.001
Commit to getting flu shot (two texts: initial text + 3 d later)	0.022	(0.004)	<0.001	<0.001
Protect others by getting a flu shot (two texts: initial text + 3 d later)	0.021	(0.004)	<0.001	<0.001
45% of Americans get the flu shot, more than in the past (two texts: initial text + 3 d later)	0.021	(0.004)	<0.001	<0.001
Receive a joke about the flu (one text)	0.021	(0.004)	<0.001	<0.001
Share a joke about the flu (one text)	0.019	(0.004)	<0.001	<0.001
People who get flu shots are less likely to get the flu (one text)	0.016	(0.004)	<0.001	<0.001
Get a flu shot to avoid getting sick (one text)	0.015	(0.004)	<0.001	<0.001
Get a flu shot to avoid getting sick and reminder of previous sickness (one text)	0.015	(0.004)	<0.001	<0.001
Think about risk of catching the flu (one text)	0.014	(0.004)	<0.001	<0.001
Do yourself a favor by getting flu shot (two texts: initial text + 2 h later)	0.014	(0.004)	<0.001	<0.001
Do others a favor by getting the flu shot (two texts: initial text + 2 h later)	0.012	(0.004)	0.002	0.003
People who get flu shots are healthier, wealthier, and more educated (one text)	0.011	(0.004)	0.004	0.004
Think about risk of catching the flu at specific locations (one text)	0.009	(0.004)	0.014	0.014
R-squared			0.0133	
Baseline vaccination rate (%)			29.4	
Observations			689,693	

Note: The above table reports the results of an OLS regression predicting whether patients in our study received a flu shot at Walmart between September 25, 2020 (when our intervention began), and December 31, 2020 (inclusive), with 22 different indicators for each of our experimental conditions as the primary predictors. The reference group is the business-as-usual control condition. The regression includes the following control variables: 1) patient age, 2) an indicator for whether a patient is male, 3) indicators for patient race/ethnicity (Black non-Hispanic, Hispanic, Asian, and other/unknown; white non-Hispanic omitted), and 4) racial composition of the patient's county (percent white, percent Black, and percent Hispanic; indicator for missing). Robust SEs accounting for heteroscedasticity in linear probability models are shown in parentheses. Adjusted *P* values accounting for multiple comparisons are calculated using the Benjamini–Hochberg method.

(*P* = 0.002) and 2) sent messages on multiple days (*P* = 0.035). Results were similar when predicting intervention efficacy between September 25, 2020, and October 31, 2020, as well as when, for robustness, we instead ran an OLS regression predicting individual vaccinations using the same control variables as in our main analysis (*SI Appendix, Tables S14–S17*).

Heterogeneity analyses did not show significant variation in treatment effects across different demographic groups (male versus female, under 65 versus 65+, and white versus nonwhite versus unknown; *SI Appendix, Tables S4–S6*).

Prediction Study Methods

To assess the ex ante predictability of this megastudy's results, we collected forecasts of different interventions' efficacy from two populations. First, in November 2020, we invited each of the scientists who designed one or more interventions in our megastudy to estimate the vaccination rates among patients in all 22 intervention conditions as well as among patients in the business-as-usual control condition. Twenty-four scientists participated (89% of those asked), including at least one representative from each design team, and these scientists made a total of 528 forecasts. In January 2021, we also recruited 406 survey respondents from Prolific to predict the vaccination rates

among patients in six different intervention conditions (independently selected randomly from the 22 interventions for each forecaster) as well as among patients in the business-as-usual control, which generated a total of 2,842 predictions. Participants from both populations were shown a realistic rendering of the messages sent in a given intervention and then asked to predict the percentage of people in that condition who would get a flu shot from Walmart pharmacy between September 25, 2020, and October 30, 2020. For more information on recruitment materials, participant demographics, and the prediction survey, refer to *SI Appendix*.

Prediction Study Results

The average predictions of scientists did not correlate with observed vaccination rates across our megastudy's 23 different experimental conditions (*N* = 23, *r* = 0.03, and *P* = 0.880). Prolific raters, in contrast, on average accurately predicted relative vaccination rates across our megastudy's conditions (*n* = 23, *r* = 0.60, and *P* = 0.003)—a marginally significant difference (Dunn and Clark's *z*-test: *P* = 0.048; Steiger's *z*-test: *P* = 0.051; and Meng et al.'s *z*-test: *P* = 0.055) (25–27). Further, the median scientist prediction of the average lift in vaccinations across our interventions was 6.2%, while the median Prolific

Text 1 (sent on 9/25/2020)

WalmartRx - Hi John! It's flu season & you can get a flu shot at Walmart. To help you remember, you'll receive another text in a few days. Flu vaccines prevent getting or spreading the flu. INFO = info, STOP = opt out.

Text 2 (sent on 9/28/2020)

WalmartRx - Remember a flu shot is waiting for you at Walmart.

Fig. 2. Text messages sent to pharmacy patients encouraging vaccination in our top-performing intervention.

respondent guess was 8.3%—remarkably close to the observed average of 8.9%.[‡] Notably, neither population correctly guessed the top-performing intervention. In fact, scientists' predictions placed it 15th out of 22, while Prolific raters' predictions placed it 16th out of 22 (*SI Appendix, Table S18*).

Discussion

The results of this megastudy suggest that pharmacies can increase flu vaccination rates by sending behaviorally informed, text-based reminders to their patients. Further, although all interventions tested increased vaccination rates, there was significant variability in their effectiveness. Attribute analyses suggest that the most-impactful interventions employed messages sent on multiple days and conveyed that a flu vaccine was already waiting for the patient. These insights from our megastudy of flu vaccinations can potentially inform efforts by pharmacies and hopefully also providers and governments around the world in the ongoing campaign to encourage full vaccination against COVID-19.

The first- and second-best-performing messages in our megastudy repeatedly reminded patients to get a vaccine and stated that a flu shot was “waiting for you.”[§] This aligns with prior research suggesting multiple reminders can help encourage healthy decisions (13). In terms of message content, communicating that a vaccine is “waiting for you” may increase the perceived value of vaccines, in accord with research on the endowment effect showing that we value things more if we feel they already belong to us (33). Further, because defaults convey implicit

[‡]To control for outliers, we calculated the median predicted average effect size across individuals in each cohort.

[§]If the best-performing messages from our test set would be costly to roll out (in dollars or other metrics), then another cheaper, high-performing intervention could be preferable to deploy (28). For most providers, we suspect that the marginal differences in the cost and complexity of deploying different messages tested in our megastudy would be vanishingly small. However, we hope that the transparency with which we've reported our results will allow policy makers interested in deploying our findings to make the right cost-benefit tradeoffs for their organizations (29–32).

recommendations, this message may imply that the pharmacy is recommending vaccination (34)—otherwise, why would they have allocated a dose to you? Relatedly, this phrasing implies that the patient already agrees that getting the vaccine is desirable. Finally, saying a vaccine is “waiting for you” may suggest that getting a vaccine will be fast and easy.

Remarkably, the three top-performing text messages in a different megastudy, which included different intervention messages encouraging patients to get vaccines at an upcoming doctor's visit, similarly conveyed to patients that a vaccine was “reserved for you.” (8) The Walmart megastudy tested a wide range of different messaging tactics designed by different researchers to explore largely different hypotheses such that the two megastudies had very limited overlap in their content. Further, our outcome variable here was electing to get a vaccine in a Walmart pharmacy, whereas the prior megastudy examined whether patients accepted a proffered vaccine in a doctor's office. Despite differences in the messenger (pharmacy versus primary care provider), the outcome (seeking a vaccine at a pharmacy versus passively accepting one in a doctor's office), and the messages actually tested, and despite analyzing our data without incorporating any priors, the results of the two megastudies converged, painting a consistent picture of what works. Together, this evidence suggests that sending reminder messages conveying that vaccines are “reserved” or “waiting” for patients is an especially effective communications strategy in the two most-common vaccination settings in the United States (16). While more research is needed to establish boundary conditions, it is notable that a follow-up study that built on the findings from these two megastudies and adapted the best-performing treatment to nudge COVID-19 vaccinations in the winter of 2021 found convergent results (9).

The scientists who designed interventions in our megastudy were not able to accurately forecast the average or relative performance of text message interventions designed to boost vaccinations. In contrast, lay survey respondents made fairly accurate predictions of both. This contradicts past research showing that experts are generally better at predicting the effectiveness of behavioral interventions than nonexperts (35). One possible explanation is that direct involvement in the design of interventions led to bias. Another possibility is that nonexperts are generally more-accurate relative forecasters in this particular field setting. Regardless, the inability of either scientists or laypeople to anticipate the top-performing intervention underscores the value of empirical testing when seeking the best policy.

The strengths of this megastudy include its massive, national sample, statistical power, an objective measure of vaccination obtained longitudinally over a 3-mo follow-up period, and simultaneous comparison of many different interventions. Several important weaknesses are also worth noting. First, we were only able to measure the impact of our interventions on vaccinations received at Walmart pharmacies. We cannot assess whether interventions increased vaccinations in other locations or, in fact, crowded out vaccinations in other settings. Past research found no meaningful evidence of crowd-in or crowd-out from reminder messages nudging vaccine adoption (4), but we cannot rule out either possibility. Another limitation is that the population studied only included patients who had received an influenza vaccine at a Walmart pharmacy in the prior flu season and agreed to receive SMS communications from Walmart pharmacy. Even so, our key findings are consistent with a different megastudy nudging vaccination at doctors' visits (8), in which results were not moderated by prior vaccination status. It would be ideal to replicate and extend this megastudy with patients who have no prior history of vaccination. Relatedly, while we did not observe meaningful heterogeneity in treatment effects by available demographics, future research exploring heterogeneous treatment effects with a richer set of individual difference variables would be valuable.

Data Availability. The experimental data analyzed in this paper were provided by Walmart. We cannot publicly post individual-level data on vaccinations that we receive from our pharmacy partner, but aggregated summary data are available on the Open Science Framework (https://osf.io/rn8tw/?view_only=546ed2d8473f4978b95948a52712a3c5) (36). Data containing individual-level health information are typically not made publicly available to protect patient privacy.

ACKNOWLEDGMENTS. Research reported in this publication was supported in part by the Bill & Melinda Gates Foundation, the Flu Lab, the Penn Center for Precision Medicine Accelerator Fund, and the Robert Wood Johnson Foundation. Support for this research was also provided in part by the AKO Foundation, John Alexander, Mark J. Leder, and Warren G. Lichtenstein. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the listed entities. We thank Aden Halpern for outstanding research assistance and Walmart for partnering with the Behavior Change for Good Initiative at the University of Pennsylvania to make this research possible.

1. H. Ritchie *et al.*, *Coronavirus (COVID-19) vaccinations – Statistics and research*. <https://ourworldindata.org/covid-vaccinations>. Accessed 19 January 2022.
2. National Governors Association, COVID-19 vaccine incentives. <https://www.nga.org/center/publications/covid-19-vaccine-incentives/>. Accessed 19 January 2022.
3. The White House, FACT SHEET: President Biden to announce national month of action to mobilize an all-of-America sprint to get more people vaccinated by July 4th. <https://www.whitehouse.gov/briefing-room/statements-releases/2021/06/02/fact-sheet-president-biden-to-announce-national-month-of-action-to-mobilize-an-all-of-america-sprint-to-get-more-people-vaccinated-by-july-4th/>. Accessed 19 January 2022.
4. K. L. Milkman, J. Beshears, J. J. Choi, D. Laibson, B. C. Madrian, Using implementation intentions prompts to enhance influenza vaccination rates. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10415–10420 (2011).
5. G. B. Chapman, M. Li, H. Colby, H. Yoon, Opting in vs opting out of influenza vaccination. *JAMA* **304**, 43–44 (2010).
6. D. Yokum, J. C. Lauffenburger, R. Ghazinoori, N. K. Choudhry, Letters designed with behavioural science increase influenza vaccination in Medicare beneficiaries. *Nat. Hum. Behav.* **2**, 743–749 (2018).
7. A. K. Regan, L. Bloomfield, I. Peters, P. V. Effler, Randomized controlled trial of text message reminders for increasing influenza vaccination. *Ann. Fam. Med.* **15**, 507–514 (2017).
8. K. L. Milkman *et al.*, A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101165118 (2021).
9. H. Dai *et al.*, Behavioural nudges increase COVID-19 vaccinations. *Nature* **597**, 404–409 (2021).
10. P. Sheeran, Intention—Behavior relations: A conceptual and empirical review. *Eur. Rev. Soc. Psychol.* **12**, 1–36 (2002).
11. M. daCosta DiBonaventura, G. B. Chapman, Moderators of the intention–behavior relationship in influenza vaccinations: Intention stability and unforeseen barriers. *Psychol. Health* **20**, 761–774 (2005).
12. E. T. Bronchetti, D. B. Huffman, E. Magenheimer, Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination. *J. Econ. Behav. Organ.* **116**, 270–291 (2015).
13. E. Yoeli *et al.*, Digital health support in treatment for tuberculosis. *N. Engl. J. Med.* **381**, 986–987 (2019).
14. CDC, Pharmacies participating in COVID-19 vaccination. <https://www.cdc.gov/vaccines/covid-19/retail-pharmacy-program/participating-pharmacies.html>. Accessed 19 January 2022.
15. K. Amin, M. Rae, S. Artiga, G. Young, G. Ramirez, Where do Americans get vaccines and how much does it cost to administer them? https://www.healthsystemtracker.org/chart-collection/where-do-americans-get-vaccines-and-how-much-does-it-cost-to-administer-them/#item-vaccinelocationandcostcollection_2. Accessed 19 January 2022.
16. M. C. Lindley *et al.*, *Early-Season Influenza Vaccination Uptake and Intent Among Adults – United States, September 2020*. Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases (NCIRD, 2020).
17. K. L. Milkman *et al.*, Megastudies improve the impact of applied behavioural science. *Nature* **600**, 478–483 (2021).
18. T. W. Cline, J. J. Kellaris, The influence of humor strength and humor—Message relatedness on ad memorability: A dual process model. *J. Advert.* **36**, 55–67 (2013).
19. C. R. Mortensen *et al.*, Trending norms: A lever for encouraging behaviors performed by the minority. *Soc. Psychol. Personal. Sci.* **10**, 201–210 (2019).
20. G. Sparkman, G. M. Walton, Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychol. Sci.* **28**, 1663–1674 (2017).
21. R. Vaidyanathan, P. Aggarwal, Using commitments to drive consistency: Enhancing the effectiveness of cause-related marketing communications. *J. Mark. Commun.* **11**, 231–246 (2005).
22. US Census Bureau, American Community Survey 5-year estimates. <https://github.com/MEDSL/2018-elections-unofficial/blob/master/election-context-2018.md>. Accessed 1 April 2021.
23. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
24. W. James, C. Stein, "Estimation with quadratic loss" in *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics Probability*, J. Neyman, Ed. (University of California Press, 1961), vol. 1, pp. 361–379.
25. O. J. Dunn, V. Clark, Correlation coefficients measured on the same individuals. *J. Am. Stat. Assoc.* **64**, 366–377 (1969).
26. J. H. Steiger, Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **87**, 245 (1980).
27. X. L. Meng, R. Rosenthal, D. B. Rubin, Comparing correlated correlation coefficients. *Psychol. Bull.* **111**, 172 (1992).
28. S. Benartzi *et al.*, Should governments invest more in nudging? *Psychol. Sci.* **28**, 1041–1055 (2017).
29. R. Berman, C. Van den Bulte, False discovery in A/B testing. *Manage. Sci.*, 10.1287/mnsc.2021.4207 (2021).
30. E. M. Feit, R. Berman, Test & roll: Profit-maximizing A/B tests. *Mark. Sci.* **38**, 1038–1058 (2019).
31. C. F. Manski, A. Tetenov, Sufficient trial size to inform clinical practice. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10518–10523 (2016).
32. J. Stoye, Minimax regret treatment choice with finite samples. *J. Econom.* **151**, 70–81 (2009).
33. D. Kahneman, J. L. Knetsch, R. H. Thaler, Anomalies: The endowment effect, loss aversion, and status quo bias. *J. Econ. Perspect.* **5**, 193–206 (1991).
34. C. R. McKenzie, M. J. Liersch, S. R. Finkelstein, Recommendations implicit in policy defaults. *Psychol. Sci.* **17**, 414–420 (2006).
35. S. DellaVigna, D. Pope, Predicting experimental results: Who knows what? *J. Polit. Econ.* **126**, 2410–2456 (2018).
36. Milkman *et al.*, Data from "A 680,000-person megastudy of nudges to encourage vaccination in pharmacies." Open Science Framework. https://osf.io/rn8tw/?view_only=546ed2d8473f4978b95948a52712a3c5. Deposited 4 August 2021.