Northwestern | INSTITUTE FOR POLICY RESEARCH

RESEARCH EXCELLENCE • POLICY IMPACT

IPR Working Paper Series

WP-25-6

Does Q&A Boost Engagement? Health Messaging Experiments in the U.S. and Ghana

Erika Kirgios
University of Chicago

Susan Athey
Stanford University

Angela Duckworth
University of Pennsylvania

Dean Karlan
Northwestern University and IPR

Michael Luca
Johns Hopkins University

Katherine Milkman
University of Pennsylvania

Molly Offer-Westort
University of Chicago

Version: February 5, 2025

Abstract

Effective information sharing is critical for the success of organizations and governments. Because information that is easy to access is more likely to be adopted, leaders often minimize friction in information delivery. However, one type of friction may increase engagement: piquing curiosity by posing relevant questions prior to sharing information. To test this, the researchers shared identical information about COVID-19 in either question-and-answer format or via direct statements across two preregistered field experiments in Ghana and Michigan (total N=49,395). Q&A-style communication increased information seeking about directly related topics (e.g., how to wear a mask properly) by 1.0 percentage-point (216%) in Ghana and by 1.1 percentage-points (19%) in Michigan (p's<0.001), and increased self-reported behavior change by 1.3 percentage-points (4%) in Michigan (p=0.002). However, sharing information in Q&A format did not increase interest in general COVID-19 information in either setting, suggesting that the impact of Q&A-style messaging on information seeking may be issue-specific. In Michigan, both Q&A-style and direct statement messaging produced less information-seeking than sending no informational messages, likely due to differential attrition: the more texts participants received, the more likely they were to opt out of receiving messages, which made it impossible for them to seek more information via text. In a follow-up implementation experiment with social media ads (a messaging strategy without attrition challenges) Q&A-style ads generated 9-11% more unique clicks to the CDC website per dollar spent than ads that directly stated information about vaccines (p<0.001). The researchers speculate that Q&A-style information delivery may stimulate curiosity, driving its benefits.

1. Introduction

Effective information sharing is critical for the success of organizations and governments alike. Policymakers frequently need to communicate with constituents about topics ranging from public health crises to environmental emergencies; managers similarly need to convey information about technologies, procedures, changing leadership, and core values with employees (Yang & Maxwell, 2011). So, leaders must carefully consider how to design communications to maximize engagement with the information that truly matters. One major hurdle communicators face is that attention—which is critical to effective information processing—is limited and highly selective (Loewenstein & Wojtowicz, 2023). People commonly overlook critical information (Pennycook et al., 2021; Handel & Schwartzstein, 2018; Pennycook & Rand, 2019; Jin et al., 2021; Jin et al., 2022), so merely making important information available does not guarantee people will attend to it (John et al., 2022; Jin et al., 2022; Dai & Luca, 2020).

A wealth of work in behavioral science suggests that to improve information's impact, decision-makers should reduce friction in the information's delivery (Handel & Schwartzstein, 2018; Bergman et al., 2020; Fishbane et al., 2020; John et al., 2022; Shah & Oppenheimer, 2008; Dai & Luca, 2020). For example, simplifying a court summons dramatically increases the likelihood that defendants will appear in court (Fishbane et al., 2020), and posting alerts of food safety violations on restaurants' Yelp pages decreases patronage of unsafe establishments (Dai & Luca, 2020). As these examples highlight, we use the term "friction" to describe features of the way information is delivered that create costs to accessing and processing that information (Handel & Schwartzstein, 2018). Naturally, increasing costs should reduce the information's impact; conversely, decreasing costs by reducing friction (e.g., by increasing the salience and clarity of information) should increase adoption of the information and follow-through on its recommendations. But could *adding* friction sometimes *increase* engagement? We propose and test a new way to effectively communicate key information: presenting it indirectly as the answer to a question. Despite adding friction, question-and-answer (Q&A) style information delivery may stimulate curiosity and boost information's perceived value, ultimately increasing its impact.

To test this, we ran two large, preregistered field experiments in which we shared vital information about the COVID-19 pandemic with residents of Ghana (throughout the country, Study 1) and the United States (in Michigan, Study 2; N = 11,585 and N = 29,810, respectively; see Table 1 for details of implementation similarities and differences across experiments). The COVID-19 pandemic provided a high-stakes context in which to test our theorizing, as governments around the world needed to encourage citizens to engage with public health information. Struggling with this challenge, the United Nations claimed that COVID-19 was not only a health crisis, but a communication crisis (Department of Global Communications, 2020). In our experiments, we randomly varied whether we (1) texted facts to people directly, making the information provided extremely easy to access, or (2) texted people *questions*, which required them to either text back a guess or wait 24 hours to receive an answer. Notably, our first strategy followed best practice: at the outset of the COVID-19 pandemic, the World Health Organization developed an SMS message library with fifty texts sharing information about COVID-19 to be sent globally, and all these messages included facts shared as direct statements (Digital Health and Innovation Team, 2020).

We find that communicating identical information in Q&A format rather than as a direct statement increased subsequent information seeking substantially (by 1 percentage-point, or 216%, in Ghana and by 1.1 percentage-points, or 19%, in Michigan, both with $p < 0.001$) about directly related topics (e.g., a Q&A communication about mask wearing increased information seeking about mask wearing). Moreover, sharing information Q&A-style increased self-reported adherence to recommended health behaviors by 1.3 percentage-points, or 4.4% ($p = 0.002$). However, communicating information in Q&A format did not boost participants' general interest in learning about COVID-19 or in sharing information about COVID-19 with others.

Relative to Michigan residents in an untreated baseline control condition, those who received information (either as direct statements or in Q&A format) were more likely to opt out of receiving messages completely. We analyze intent-to-treat effects from our information provision treatments, assuming the most "pessimistic" responses for unresponsive participants. Given this analysis strategy and differential opt-out rates across conditions, sharing information in either a direct statement or Q&A format via text message mechanically appears ineffective as a means of boosting either information seeking or adherence to recommended health behaviors. Importantly, however, leaders who urgently need to communicate information (during a public health emergency or at any other time of crisis) are interested in the best way to communicate that information. They are likely to be less concerned with the fact that sending more text communications increases opt-out rates, which biases measurement of behavior change collected via subsequent (unsendable) texts. Therefore, the most relevant comparison in this study (both practically and theoretically) is between the two information sharing conditions. Theoretically, this comparison helps us understand which is the stronger driver of information engagement: ease of access or curiosity.

However, to validate the relevance of our experimental results for practitioners who may not want or need to share information via text message, we conducted an implementation experiment (Study 3) on social media. Specifically, we tested whether a Facebook ad that posed a question about vaccines would prove more cost-effective than an ad that directly shared the corresponding fact about vaccines. Using a fixed daily ad spend and optimizing cost-per-click, we find that Q&A-style ads generated 9.1% more unique clicks to the CDC vaccine webpage than ads that shared facts directly ($p < 0.001$). Our experimental results confirm that despite adding friction (by requiring people to exert effort or endure a delay before accessing key information), posing a question rather than providing the answer alone can increase information engagement.

Generally, there is good reason to believe increasing friction will be harmful rather than helpful. People often ignore or simply do not notice valuable information that must be acquired with even minimal effort. For example, one study of parents found that only 11% texted “start” to receive updates about their children’s progress in school when offered the opportunity to do so. This information neglect was costly, increasing the likelihood that their children failed a class by 9% relative to the children of parents who were automatically sent progress updates (Bergman et al., 2020). Relatedly, simplifying the language in court summons to make important information about when, where, and why to appear in court more salient reduced failures to appear in court by 13-21% (Fishbane et al., 2020). Overall, past research has demonstrated that making information easier to access and process can improve its influence, often with meaningful welfare benefits.

However, reducing friction may not always be the ideal path for maximizing attention. Friction that builds desire to resolve uncertainty may improve information's impact. Recent work suggests that persuasive appeals interrupted by a brief pause (e.g., a spinning wheel indicating that a video needed time to finish loading) were more convincing than uninterrupted appeals because the waiting period increased people's interest in what came next (Kupor & Tormala, 2015). Similarly, we suggest that sharing information only after "teasing" people with a question may stimulate curiosity and ultimately increase engagement, even though it adds friction to the learning process.

Curiosity is triggered when a piece of missing knowledge or a space between what someone knows and what they want to know (an "information gap") is made salient (Loewenstein, 1994).[1] Because people are motivated to resolve uncertainty, the feeling of not knowing may increase the perceived value of information that could close the gap, and thus the resources they are willing to devote to acquiring it (Loewenstein, 1994; Hsee & Ruan, 2016; Shen et al., 2015; Polman et al.,, 2022; Horn et al., 2024). People are willing to undergo electric shocks, forgo tempting foods, and walk several flights of stairs to resolve their curiosity (Hsee & Ruan, 2016; Polman et al., 2022). Moreover, the feeling of closing information gaps can be quite rewarding, like sating hunger or thirst (Kang et al., 2009; Wojtowicz & Loewenstein, 2020). As a result, people report higher enjoyment during the information acquisition process when their curiosity is aroused. Since boosting enjoyment of a task increases persistence, curiosity may similarly increase information-seeking behavior (Woolley & Fishbach, 2016; Ruan et al., 2018; Horn et al., 2024). Together, this work suggests that people who receive information in a Q&A

[1] Extant work often uses questions to stimulate curiosity, suggesting that our Q&A condition will also activate curiosity (e.g., Horn et al., 2024; Polman et al., 2022; Loewenstein et al., 1998). We rely on this prior work to establish that curiosity will be activated by Q&A-style messaging, as our focus is on three field tests of our intervention's effectiveness rather than on measuring curiosity or other intermediate psychological outcomes in the lab.

format rather than in the form of direct statements may be more engaged and perceive the information they obtain to be more valuable, even though they are required to exert more effort to get it.

Anyone aiming to widely communicate important information (from governments to managers) is battling for attention. Our work suggests that one way to increase information engagement is to boost curiosity by first asking questions, then providing answers. Even if this strategy creates some friction, people are more likely to engage with and utilize critical information after being prompted with a related question.

2. Study 1: Field Experiment in Ghana

In a preregistered (https://aspredicted.org/xp8c-v66w.pdf) field experiment, we test whether texting people important health information in a Q&A format encourages more information seeking and information sharing than delivering the same information directly. De-identified data, analysis code, and a Supplementary Materials file with complete study stimuli and further details about study implementation and results for Studies 1-3 are available on OSF at https://osf.io/5ab26/.[2]

2.1 Methods

We recruited Ghanaian adults using random digit dialing via phone survey and Interactive Voice Response (IVR) phone calls between May 22, 2020 and December 1, 2020. At the end of the IVR call, we asked if they wanted to opt in to receive text messages about COVID-19.[3] Out of 204,617 people contacted, 21,403 (10.5%) opted into the study and 11,585 were successfully randomly assigned to experimental conditions.[4]

---

[2] This research was approved by the Institutional Review Board at the first author's institution and complies with all relevant ethical regulations. We received a waiver of informed consent for Study 3, and informed consent was obtained from all study participants in Studies 1 and 2. Participants were not compensated.

[3] All recruiting was conducted in English, providing an implicit screening for English literacy, and all intervention messages were sent in English as well. Overall, about 67% of the Ghanaian population can read and write in English (Ghana Statistical Service, 2013).

[4] Our preregistered sample size was 21,403. At the end of our intervention, we learned of an error that occurred when our implementation partner downloaded the file with participant random assignment. Before the experiment began, 4,910 participants in one experimental condition were overwritten with 4,908 participants in the other condition. So, 4,910 participants never received a text message while 4,908 received messages from both experimental conditions. Consistent with our preregistered plan for dealing with glitches in random assignment, we excluded this set of 9,818 participants, leaving us with our final sample size of 11,585. Importantly, assignment to

Despite opting in at different times, all participants began receiving intervention text messages on December 10, 2020 and stopped receiving messages after February 1, 2021. During this period, we sent participants an average of 3.75 SMS messages per week. None of the 11,585 participants opted out of the intervention at any point during the eight-week intervention period,[5] perhaps because opting out required sending a text message, which typically costs $0.11 in Ghana. Although we covered all messaging costs for participants, many were still concerned about reply costs (see "Additional Implementation Details" in the Supplement).

We randomly assigned all 11,585 participants to one of two experimental conditions: the *Q&A* condition or the *direct statement* condition. Participants in both conditions received messages containing information about COVID-19 in eight separate information-sharing message sequences (see Figure 1 for a visualization of the timing of all text messages, Supplement Table 22 for the contents of each text message, and "Additional Implementation Details" in the Supplement for more information about how the messages were crafted and adapted to the local environment). In the *direct statement* condition (N = 5,741), information was shared as a statement of fact (e.g., "COVID TEXT: Masks aren't enough to protect you on their own: Even if you and another person are both wearing masks, you need to stay 1m apart to be safe."). In the *Q&A* condition (N = 5,844), information was shared by first asking participants a question (e.g., "COVID TEXT: If you and another person are both wearing masks, do you still need to stay 1m apart to stay safe? Reply 1 for Yes, 2 for No"). If participants replied to the question, they received the correct answer immediately (e.g., if a participant replied incorrectly by texting "2," they would then get a text that said: "Actually, masks aren't enough to protect you on their own: Even if you and another person are both wearing masks, you need to stay 1m apart to be safe." If, instead, the participant replied correctly, the fact would instead be prefaced by "That's right!"). If they did not reply,

---

each condition remained random: the effect of these exclusions was to remove a random subset of participants of about equal size from each of our experimental conditions.

[5] The intervention period was originally intended to be four weeks. However, due to delays in the phone lines in Ghana, limitations in the speed with which our implementation partner could send text messages, and glitches that occurred during the intervention, the intervention lasted nearly eight weeks instead (see Figure 1 for a timeline).

they received the answer after 24 hours regardless.[6] In sum, no matter how participants in the *Q&A* condition responded, they received the same information as participants in the *direct statement* condition—the key difference was whether a question preceded this information and therefore some effort (or time delay) was required to obtain it.[7] Note that this design implies that participants in the *Q&A* condition also always received one additional text in information-sharing message sequences relative to participants in the *direct statement* condition, because they received a text containing a question before receiving information.

We collected two primary types of dependent variables during the intervention: (1) measures of information seeking and (2) measures of information sharing.[8]

Measures of Information Seeking. Participants were given five opportunities to seek more information about COVID-19 over the eight-week intervention period. Specifically, participants were twice prompted to text "1" to access more free information about COVID-19 immediately after being

---

[6] For each question sent in the *Q&A* condition, fewer than 2% of participants responded. So, most participants waited 24 hours to receive the information.

[7] We conducted a preregistered post-test with 904 participants on Prolific to assess whether delivering information in Q&A style, as in this experiment, really does increase friction (i.e., the cost of accessing and processing the information; Handel & Schwartzstein, 2018) relative to delivering information in direct statements. Post-test participants learned five facts about mpox, formerly known as monkeypox. They were randomly assigned to either read five direct statements of fact about mpox, or to receive the same information in a Q&A format. Specifically, to mimic the design of Study 1, participants assigned to learn facts in Q&A format first read a question about the fact and then either guessed the answer by typing into a text box or waited 15 seconds to receive the answer if they chose not to guess. After learning the five facts, all participants responded to a fluency scale (to measure ease of processing; adapted from Kostyk, Leonhardt, & Niculescu, 2019) and a three-item, face-valid scale intended to measure ease of access (specifically: "It was easy to learn about mpox through this health messaging program", "It was straightforward to access information about mpox through this health messaging program", and "Learning about mpox through this health messaging program was annoying", reverse scored). We found that participants perceived the information shared to be directionally more difficult to process ($p = 0.08$) and significantly more difficult to access ($p < 0.001$) when it was delivered in Q&A-style rather than via direct statements, providing suggestive evidence that sharing information only after "teasing" people with a question does increase friction. See Supplementary Study 1 in the Supplement for more details.

[8] We also collected a third category of dependent variables: self-reported behaviors. We collected four separate self-reports of adherence to recommended health behaviors (e.g., whether participants washed their hands after their last outing). However, response rates to these self-reported behavior questions were at or below 1.5%. While non-responses are meaningful for information seeking and information sharing, they are uninformative for self-reported behaviors (for the ~99% of participants who did not respond to these messages, we cannot infer whether or not they washed their hands with reasonable levels of accuracy; meanwhile, if someone does not respond to an offer for more information, we know they did not engage in information-seeking behavior). As a result, we cannot make strong inferences from our data for this group of dependent variables. We include a further description of our methods and results for self-reported behavior change in the Supplement.

asked to self-report their behaviors (see Figure 1). Additionally, on three separate occasions, participants were prompted to text "1" to learn more about a specific COVID-19-related topic immediately after an information-sharing message sequence (the topics were: how to wear a mask properly, how to safely run errands, and which populations are at the highest risk for the disease). Both types of information seeking opportunities (general and specific) were presented as statements (e.g., "TEXT 1 if you'd like a few tips for how to stay safe when you go out") rather than as questions (see Supplement Table 22). Our preregistered dependent variables are (1) an information-seeking composite representing the total number of times (out of five) a participant sought more information and (2) indicators for whether participants sought information at each of the five information-seeking opportunities we presented.

Measure of Information Sharing. Halfway through the intervention, participants were offered the opportunity to share information about COVID-19 by texting us their friends' and family members' phone numbers. We let them know that we might text or call the numbers they provided. Our preregistered dependent variable of interest was the number of phone numbers participants shared.

2.2 Results

Analytic Plan. Following our preregistration, we relied on an ordinary least squares (OLS) regression with robust standard errors to identify the effects of assignment to the *Q&A* condition on each of our dependent variables of interest:

$$Y_i = \beta Q_i + \gamma T_i + \varepsilon_i$$

Y_i is the dependent variable of interest measured for participant *i*, Q_i is an indicator for whether participant *i* was assigned to the *Q&A* condition, and T_i is a continuous control for the number of days prior to the intervention start date that participant *i* was recruited via IVR.[9] Because we did not collect any

[9] As a robustness check, we also ran two sample t-tests to compare our continuous dependent variables across conditions and proportions tests to compare our binary dependent variables across conditions, and our results are generally robust to either analysis strategy (see Supplement Table 1). The time between recruitment and the intervention start date varied widely across participants. It took us several months to recruit our participant pool, but regardless of when they were first contacted via IVR, all participants started the intervention on the same day due to technology constraints faced by our implementation partner in Ghana. As might be expected, the time lag between recruitment and the start of the intervention was significantly negatively related to likelihood of responding to messages (see Table 2). For summary statistics describing the time between recruitment and intervention delivery, see Supplement Table 1.

information about participants, we could not control for any participant characteristics, and we could not conduct balance checks to ensure randomization was successful across conditions. However, there was balance across conditions on our one control variable, T_i (p = 0.25). All analyses were conducted intent-to-treat.

Information Seeking. On average, participants sought more information 0.07 times out of 5 (SD = 0.34), so we had 95% power to detect a 0.02 increase in information seeking across conditions.

Participants assigned to the *Q&A* condition took advantage of 0.02 more information seeking opportunities than those in the *direct statement* condition (95% CI: [0.01, 0.04], p < 0.001; raw estimates: 0.08 out of 5 in the *Q&A* condition vs. 0.05 out of 5 in the *direct statement* condition; see Table 2, Model 1 for full regression results). This represents a 43% boost in information seeking in the *Q&A* condition relative to the *direct statement* condition.

Examining each information-seeking opportunity separately, we find that participants were more likely to seek information about two of the three specific topics offered in the *Q&A* condition than in the *direct statement* condition (participants were five times as likely to seek more information about mask wearing and twice as likely to seek more information about how to leave home safely in the *Q&A* condition, both with p < 0.01) and marginally more likely to seek information about the third topic (p = 0.09); see Figure 2 and Supplement Table 1. However, we did not observe this effect when we offered to share more general information about COVID-19. At the start of the 8-week intervention, after receiving only two information-sharing message sequences, participants were just as likely to ask for more general information in the *Q&A* condition as in the *direct statement* condition (p = 0.52). When we again offered them an opportunity to learn more general information about COVID-19 at the end of the eight-week intervention period, participants were significantly less likely to ask for it in the *Q&A* condition than in the *direct statement* condition (p = 0.04).[10]

[10] This reversal in our effect on day 52 of the intervention may be the result of an implementation glitch that occurred on day 41. Specifically, our implementation partner launched an IVR survey to attempt to diagnose the low response rates to our text messages (see "Additional Implementation Details" in the Supplement for the results of this IVR). Unfortunately, they accidentally launched the IVR survey to the *Q&A* condition only. To attempt to fix

In exploratory analyses, we created two separate composite variables. The first composite measured people's overall tendency to seek topic-specific information about COVID-19 and the second measured people's overall tendency to seek general information about COVID-19. We relied on the same OLS regression analysis strategy deployed above to compare likelihood of seeking more topic-specific (and general) information across conditions.

Our composite analyses demonstrate that on average, participants assigned to the *Q&A* condition took advantage of 0.03 more opportunities to seek more information about specific topics than participants in the *direct statement* condition (95% CI: [0.02, 0.04], $p < 0.001$; see Table 2, Model 2 for full regression results), representing a 216% increase in topic-specific information seeking (raw estimates: 0.04 out of 3 in the *Q&A* condition, 0.01 out of 3 in the *direct statement* condition). Moreover, the proportion of participants seeking specific information at least once nearly tripled across conditions, from 1.1% in the *direct statement* condition to 3.2% in the *Q&A* condition (two-tailed proportions test $z = 7.61$, 95% CI: [0.02, 0.03], $p < 0.001$). However, participants in the *Q&A* condition were no more or less likely to seek access to general information about COVID-19 than those in the *direct statement* condition ($\beta = -0.004$, 95% CI: [-0.01, 0.003], $p = 0.26$; see Table 2, Model 3 for full regression results).[11] Although we did not expect to find differences in the effects of Q&A-style communication on specific vs. general information seeking opportunities a priori, we explore this pattern again in Study 2 and speculate about why this divergence may arise in our General Discussion.

---

this issue, they launched a new IVR survey to participants in the *direct statement* condition on day 48 of the intervention. This inconsistency between conditions could have contributed to our results on day 52, as participants in the *direct statement* condition received the IVR survey more recently and may have therefore been more likely to answer when they received further texts from us. This is the only dependent variable that was collected after the glitch, so no other data should be affected by the implementation error and our results are robust to the exclusion of this data. See Supplement Table 2 for regression results for each information-seeking opportunity and Supplement Table 3 for results using logistic regression rather than OLS regression.

[11] Note that if we exclude the data collected after the glitch, only the general information composite is affected. In that case, we only have one general information seeking opportunity in the data (on day 14 of the intervention), and we again find that participants in the Q&A condition were no more or less likely to seek general information than those in the direct statement condition ($\beta = 0.001$, 95% CI: [-0.003, 0.01], $p = .52$; raw estimates = 0.02 out of one in the Q&A condition, 0.01 out of one in the direct statement condition).

Information Sharing. On average, for every 1,000 participants invited to share the phone numbers of people they would like to keep informed about COVID-19, 2 phone numbers were shared in the *Q&A* condition and 1 was shared in the *direct statement* condition (overall SD = 0.04). Given the low baseline mean of information sharing, we had 95% power to detect an effect size d = 0.07, or a 0.003 increase in the average phone numbers shared across conditions—a larger effect than we found. Indeed, the rate of sharing did not differ significantly across conditions (β = 0.001, 95% CI: [-0.000, 0.003], p = 0.10).

2.3 Discussion

The results of Study 1 suggest that delivering health information in Q&A format encourages more information seeking than delivering the same information as a direct statement. These results provide some evidence that sharing information in Q&A format increases engagement with the information, perhaps by stimulating participants' curiosity.

A potential limitation of Study 1 is that response rates to our messages were extremely low (i.e., between 0.5% and 3.5%). Low response rates may have been driven by a few factors: concerns about messaging costs (although costs were covered, participants may have forgotten or mistrusted this fact), the legibility of our messages (all our messages were sent in English, so some participants in our pool may not have been able to understand them), or technical issues (there were significant delays in the phone lines in Ghana during our experiment, so participants may not have received some messages). Given these low response rates, the changes in information seeking we document across conditions are necessarily small in magnitude. However, it is notable that despite the impediments to responding that our participants faced, sending messages in Q&A format nearly tripled their willingness to actively seek more information about COVID-19. Overall, about 123 more participants in the *Q&A* condition than in the *direct statement* condition sought more information about COVID-19 at least once (in a total sample of 5,844), which is arguably a meaningful impact from a relatively low-cost intervention. To concretize the potential impact of such an intervention, if Study 1 were expanded to include the entire population of Ghana aged 15+, the 2.1 percentage-point increase in the number of people who sought more information at least once in the *Q&A* condition would lead to over 400,000 more citizens actively engaged in learning more about public health.

3. Study 2: Field Experiment in Michigan

In Study 2, we set out to replicate and extend the effects of Study 1 by measuring not only information seeking and information sharing in a new context (in the United States) but also self-reported behavior change. Again, we sent text messages with public health information in a Q&A versus direct statement format. In addition, we added an inactive, *no-treatment control* condition to our study design. Participants in the *no-treatment control* condition did not receive the information-sharing message sequences and instead only received messages intended to capture our dependent variables of interest. This study was preregistered on AsPredicted.org (https://aspredicted.org/zrpp-9hqp.pdf).

3.1 Methods

We partnered with the Michigan Department of Health and Human Services (MDHHS) to send text messages with information about COVID-19 to Michigan residents who had previously consented to receive texts from government programs. Between November 12, 2020 and December 30, 2020, we sent recruitment text messages to ~800,000 phone numbers the MDHHS provided asking recipients if they were willing to receive text messages with information about COVID-19 over the next four weeks. The 29,810 participants who said yes began receiving text messages at 9am the day after they opted in and continued to receive messages throughout a four-week intervention period, unless they opted out of receiving further messages during the intervention. Thus, participants in Study 2 (unlike those in Study 1) did not face a delay between opting into the experiment and beginning the intervention; instead, we allowed for a rolling start date across participants based on when they opted in.

Participants who opted in to receive COVID-19 texts were randomly assigned to one of three conditions with weighted random assignment. Specifically, participants were assigned to a *Q&A* condition, a *direct statement* condition, or a *no-treatment control* condition in a 2:2:1 ratio. Thus, about half as many participants were assigned to the *no-treatment control* condition (N = 5,930) as either of the other conditions (*Q&A* condition N = 11,889; *direct statement* condition N = 11,991). We prioritized recruiting participants to the *Q&A* and *direct statement* conditions both because the MDHHS wanted to maximize the number of treated individuals and because our primary aim was to replicate and build on the results of Study 1.

The *Q&A* and *direct statement* conditions were similar to the experimental conditions in Study 1: twice a week for four weeks, participants received information about COVID-19 as either a fact in the *direct statement* condition or as a question in the *Q&A* condition. We call each of these eight instances an information-sharing message sequence, and the time between message sequences was constant across participants (see Figure 3 for a visualization of the timing of all messages sent in Study 2). Ultimately, the same information about COVID-19 was shared across the two conditions—all that varied was whether a question preceded the information (and, as a result, the total number of texts sent across conditions). As in Study 1, participants were given 24 hours to answer questions in the *Q&A* condition and received the answer as soon as they engaged, or after 24 hours if they failed to reply.[12] Figure 4 depicts an example information-sharing text sequence from Study 2.

In the *no-treatment control* condition, participants received no information. Instead, participants were only contacted when we collected data on our study's dependent variables. Thus, participants in the *Q&A* and *direct statement* conditions received eight information-sharing text message sequences during the intervention, while those in the *no-treatment control* condition received none (see Figure 3). We collected three types of dependent variables at regular intervals during the four-week intervention period: (1) measures of (self-reported) actions people took to prevent the spread of COVID-19, (2) measures of information seeking, and (3) measures of information sharing.[13]

Measures of self-reported actions people took to prevent the spread of COVID-19. Once a week, participants were asked to answer a question about whether they engaged in behaviors to prevent the spread of COVID-19. Specifically, we asked participants whether they wore a mask when leaving home,

[12] Participants interacted with the questions at a much higher rate in Study 2 than Study 1, with about 30-40% of participants responding and the other 60-70% receiving an answer 24 hours later.
[13] Due to our higher response rates in Study 2 relative to Study 1, we were able to confidently move forward with our preregistered analyses for self-reported behavior change.

attended social gatherings, washed their hands after their last outing, and engaged in physical contact with anyone outside their household (see Supplement Table 23 for the content and timing of each text message). When relevant, responses were reverse-scored such that higher values indicated higher levels of self-reported adherence to recommended health behaviors. Our preregistered dependent variable was a behavior change composite score created by z-scoring each of the four self-reported behaviors and averaging them together. We also created an alternative, non-preregistered behavior change composite score by counting the number of times (out of four) participants reported fully adhering to public health guidelines. We present the effect of the experimental conditions on each behavior individually as well as on these two composite scores.

Information Seeking. Participants were given five opportunities to seek more information about COVID-19 during the intervention. Immediately after the first and fourth self-reported behavior variables were collected (see Figure 3), we asked participants to text "1" for more information about COVID-19. Meanwhile, on days 10, 18, and 23 of the intervention, we asked participants to text "1" for specific information about the protective power of masks, how to leave home safely, and at-risk populations, respectively. These three specific information-seeking opportunities immediately followed information-sharing message sequences.[14] Our preregistered dependent variables were (1) an information seeking composite variable representing the number of times (out of five) participants asked for more information and (2) indicators for whether participants sought more information each of the five times it was offered.

Information Sharing. On day 12 of the intervention, we texted an invitation to participants to reply with the phone numbers of friends and family members with whom they would like us to share information about COVID-19. Our preregistered dependent variable was the number of phone numbers participants shared.

3.2 Results

[14] As in Study 1, all information seeking opportunities (both specific and general) were expressed as statements rather than questions (see Supplement Table 23).

Differential attrition. In the *direct statement* condition, 12% of participants opted out of receiving further messages at some point during the four-week intervention period, while 18% of participants in the *Q&A* condition opted out (two-tailed proportion test p < 0.001). Opt-out rates were lowest in the *no-treatment control* condition, at 10% (significantly lower than the level in the *direct statement* condition, two-tailed proportion test p = 0.01). We conducted all analyses intent-to-treat, so participants who opted out of the intervention remained in our analyses. For our self-reported behavior change measures, we used several different methods to handle missing data[15] (our preregistered method for handling missing data was to replace it with the "worst" possible response—non-adherence to recommended behaviors; we also evaluated our results by replacing missing data with the "best" possible response and with the average response).[16] The results of these robustness checks are presented in Supplement Tables 6 and 14.

3.2.1 Comparing Q&A and Direct Statement information delivery. We first present results comparing the effects of the *Q&A* condition to the *direct statement* condition. Summary statistics across analyses are reported in Supplement Table 4.

Analytic plan. Following our preregistration, we relied on an OLS regression with robust standard errors to identify the effects of assignment to the *Q&A* condition relative to the *direct statement* condition on each of our dependent variables of interest:

$$Y_i = \beta Q_i + \alpha C_i + \gamma T_i + \varepsilon_i$$

Y_i is the dependent variable of interest measured for participant *i,* Q_i is an indicator for whether participant *i* was assigned to the *Q&A* condition, C_i is an indicator for whether participant *i* was assigned to the *no-treatment control* condition, and T_i is a continuous control for the date participants opted into the study. Because we did not collect any information about participants, we could neither control for any

---

[15] Response rates to our self-reported behavior questions across conditions are compared in our Supplementary Materials in Supplement Table 4 and are visualized in Figure 5.

[16] Missing data does not require interpolation for our other dependent variables capturing information seeking and information sharing because participants were only asked to reply to texts if they wanted to actively seek or share more information. A missing value is informative in that case: Participants did *not* want more information about COVID-19 from us, and they did *not* want us to share information about COVID-19 with their friends or family members. Meanwhile, for our self-reported behaviors, missing values are not informative: there is no way to know whether a participant who did not respond to our message did or did not wash their hands after returning home from their last outing, for example.

participant characteristics nor conduct balance checks to test for successful random assignment. However, there was balance across conditions on our one control variable, T_i (p = 0.30, see Supplement Table 4 for summary statistics).

Self-reported behaviors. Examining our primary z-scored behavior change composite, we had 95% power to detect a 0.02 standard deviation increase in self-reported behaviors to stem the spread of COVID-19. We found that assignment to the *Q&A* condition led to a 0.04 standard deviation increase in self-reported adherence to recommended health behaviors relative to the *direct statement* condition (95% CI: [0.02, 0.05], p < 0.001; see Table 3, Model 1 for full regression results).[17] This result remained consistent when we focused on our summed composite score; participants in the *Q&A* condition reported adhering to public health guidelines 0.05 more times out of 4 relative to participants in the *direct statement* condition (95% CI: [0.02, 0.09], p = 0.002; raw estimates: 1.23 times out of 4 in the *Q&A* condition, 1.18 times out of 4 in the *direct statement* condition; see Table 3, Model 2 for full regression results). This represents a regression-estimated 4.4% increase in self-reported behavior change in the *Q&A* condition.

Notably, when we decompose our composite variables into weekly measures of behavior change, we identify a substantial shift in estimated treatment effects over time (see Figure 5). In the first week of the study, participants in the *Q&A* condition reported lower adherence to recommended health behaviors as compared to participants in the *direct statement* condition (p = 0.02). In week two of the intervention, there was no significant difference between conditions (p = 0.61). In weeks three and four of the intervention, however, participants reported significantly more adherence to recommended health

[17] As preregistered, missing values are treated as the "worst" possible response (e.g., missing values were replaced with a "0" for week 3, as though participants had indicated that they did not wash their hands). Given the response rates in this experiment (depending on the week, between 25-50% of participants responded to our self-reported behavior questions, see Figure 5 and Additional Analyses section in the Supplement for more details), we also present robustness checks where we handle non-responses differently in Supplement Table 6. In general, our results are directionally robust to different interpolation strategies, with the exception of treating missing values as the "best" possible response.

behaviors in the *Q&A* condition than in the *direct statement* condition (both p's < 0.001). See Table 4, Models 1-4 for full regression results.

Information seeking. There was no significant difference in overall information seeking across the *Q&A* and *direct statement* conditions. On average, participants in the *Q&A* and *direct statement* conditions took advantage of 0.36 out of five information seeking opportunities (SD = 0.81), giving us 95% power to detect a 0.02 change in willingness to seek more information. Participants assigned to the *Q&A* condition took advantage of 0.01 more information seeking opportunities (out of five) relative to those in the *direct statement* condition (95% CI: [-0.01, 0.03], $p = .38$; raw estimates: 0.36 out of 5 in the *Q&A* condition vs. 0.35 out of 5 in the *direct statement* condition; see Table 5, Model 1).

However, as in Study 1, the pattern of information seeking results differed depending on whether participants were offered general information about COVID-19 or specific information about a given COVID-19 topic, such as masking. As shown in Supplement Figure 1 and Supplement Table 8, participants were significantly more likely to ask for specific information about COVID-19 topics in the *Q&A* condition than in the *direct statement* condition on two out of three occasions. Specifically, participants in the *Q&A* condition were significantly more likely to seek more information about mask wearing (day 10, $p < 0.001$) and about how to leave home safely (day 18, $p < 0.001$), but not about at-risk populations (day 23, $p = 0.86$). Meanwhile, participants consistently sought less access to general information about COVID-19 in the *Q&A* condition than in the *direct statement* condition (on day 5, $p < 0.001$ and on day 28, $p = 0.01$). Robustness checks using logistic regression rather than OLS regression are presented in Supplement Table 9.

As in Study 1, we created two exploratory composite variables to measure overall rates of seeking topic-specific and general information across conditions. Participants in the *Q&A* condition took advantage of 0.03 more topic-specific information seeking opportunities (out of 3) than those in the *direct statement* condition (95% CI: [0.02, 0.05], $p < 0.001$; see Table 5, Model 2), representing a regression-estimated 19% increase in topic-specific information seeking. Specifically, participants took advantage of an average of 0.20 out of 3 topic-specific information seeking opportunities in the *Q&A* condition, relative to 0.17 out of 3 in the *direct statement* condition. Moreover, the proportion of participants seeking specific information at least once increased from 12.4% in the *direct statement* condition to 14.8% in the *Q&A* condition (two-tailed proportions test z = 5.49, 95% CI: [0.02, 0.03], p < 0.001). Meanwhile, participants in the *Q&A* condition took advantage of 0.02 fewer general information seeking opportunities (out of 2) than those in the *direct statement* condition (95% CI: [-0.03, -0.01], p < 0.001; see Table 5, Model 3), representing a regression-estimated 12% decrease in general information seeking. Specifically, participants took advantage of an average of 0.16 out of 2 general information seeking opportunities in the *Q&A* condition, relative to 0.18 out of 2 in the *direct statement* condition.[18]

Information sharing. On average, for every 1,000 participants offered the opportunity to provide phone numbers of people they would like to keep informed about COVID-19 there were 30 phone numbers shared in the *Q&A* condition and 31 in the *direct statement* condition (β = -0.00, 95% CI: [-0.01, 0.004], p = 0.68; see Table 5, Model 4).

3.2.2 Comparing information provision and no-treatment control. We next compare our *no-treatment control* condition with our two *information provision* conditions (the *Q&A* and *direct statement* conditions), allowing us to assess the overall impact of delivering health information via text message. Summary statistics across analyses are presented in Supplement Table 10.

Analytic plan. As preregistered, we pooled the *Q&A* and *direct statement* conditions to form an *information provision* condition. Following our preregistration, we relied on an OLS regression with

[18] There are both theoretical and mechanical reasons why there might be a decreased interest in general information seeking in the *Q&A* condition. Participants in the *Q&A* condition may have felt better informed because the interactive back-and-forth intervention messaging style kept them engaged, so they may have felt less of a need to access general resources with additional COVID-19 information than those in the *direct statement* condition. Or, enjoyment of the interactive style of knowledge transfer in the *Q&A* condition may have reduced their interest in accessing a resource with information presented in non-interactive style. However, there is also a plausible mechanical explanation: this effect may be an artifact of our intent-to-treat strategy and the differential attrition we document across conditions. Because fewer participants stuck with the intervention in the *Q&A* condition, fewer were available to indicate interest in the information. Meanwhile, in the case of specific information seeking, interest was so much higher in the *Q&A* condition that the difference shone through despite the differential attrition and our intent-to-treat analysis strategy.

robust standard errors to identify the effects of assignment to the *information provision* condition relative to the *no-treatment control* condition on each of our dependent variables of interest:

$$Y_i = \beta I_i + \gamma T_i + \varepsilon_i$$

Y_i is the dependent variable of interest measured for participant *i*, I_i is an indicator for whether participant *i* was assigned to the pooled *information provision* condition[19], and T_i is a continuous control for the date participants opted into the study.[20] Because we did not collect any information about participants, we could not control for any participant characteristics, and we could not conduct balance checks to ensure randomization was successful across conditions. However, there was balance across conditions on our one control variable, T_i (p = 0.22). All analyses were conducted intent-to-treat.

Self-reported health behaviors. Because participants were more likely to drop out of our study the more frequently we texted them (see Figure 6), and because we impute the worst possible outcomes for missing data, we mechanically saw worse outcomes in the pooled *information provision* condition. Specifically, we found that participants in the *information provision* condition reported a 0.11 standard deviation decrease in adherence to recommended health behaviors as compared to the *no-treatment control* condition (95% CI: [-0.13, 0.09], p < 0.001; see Supplement Table 11), using our primary z-scored behavior change composite. Further, examining our alternative binary composite score, participants in the pooled *information provision* condition reported adhering to recommended health behaviors 0.19 fewer times (out of 4) than participants in the *no-treatment control* condition (95% CI: [-0.23, -0.16], p < 0.001; see Supplement Table 11), representing a regression-estimated 14% decrease in adherence to public health recommendations. However, these results were not robust to different methods of handling missing data (see Supplement Table 14), which suggests it is unlikely that information provision changes health behavior in harmful ways. Instead, our findings robustly show that information

[19] As an exploratory analysis, we tested whether our results differed when we considered each of the information provision conditions (i.e., the *Q&A* and *direct statement* conditions) separately and compared them to the *no-treatment control* condition. Whether we separate the conditions or pool them, our results remain consistent, so we present the preregistered pooled analyses here.

[20] We also present the results of t-tests and proportions tests comparing results across conditions in Supplement Table 10, and for binary dependent variables, we present the results of logistic regression analyses in Supplement Tables 13 and 18.

provision changes responsiveness to text messages about behavior. Full regression results for each of the four self-reported behaviors are in Supplement Table 12.

Information seeking. Participants in the pooled *information provision* condition took advantage of 0.18 fewer information seeking opportunities (out of 5) than those in the *no-treatment* control condition (95% CI: [-0.12, -0.10], $p < 0.001$; raw estimates = 0.36 out of 5 in the *information provision* condition, 0.53 out of 5 in the *no-treatment control* condition; see Supplement Table 16 for full regression results). Indeed, participants were significantly more likely to seek additional information in the *no-treatment control* condition on four of the five information-seeking opportunities (all with $p < 0.001$ with the exception of the information seeking opportunity on day 28, $p = 0.34$; see Supplement Table 17 for full regression results and Supplement Table 18 for robustness checks using logistic regression).

Information sharing. On average, for every 1,000 participants offered the opportunity to provide phone numbers of people they would like to keep informed about COVID-19, 34 phone numbers were shared in the *no-treatment control* condition while 30 were shared in the *information provision* conditions (β = -0.01, 95% CI: [-0.01, -0.00], $p = 0.04$; see Supplement Table 16).

3.3 Discussion

Study 2 replicates the key results from Study 1 with a large sample of Michigan residents: sharing information in a Q&A format inspires more information seeking about specific COVID-19-related topics than sharing information via direct statements. However, in Study 2 we find that the effect is reversed when participants are given an opportunity to seek information about general resources related to COVID-19. We also find evidence that Q&A-style information communication may increase participants' adherence to recommended health behaviors relative to directly communicating the same facts. This is a conservative test of the benefits of sharing information in Q&A format given that all our analyses were conducted intent-to-treat with the "worst" possible responses imputed for missing values, and more participants dropped out of the intervention in the *Q&A* condition than in the *direct statement* condition. Although this is promising evidence that posing questions about health information increases willingness to act, self-report measures are noisy proxies for actual behavior.

Study 2 also included an untreated baseline control condition, and participants assigned to that condition were less likely to opt out of receiving messages they had elected to receive about public health than those assigned to either of our information provision conditions. This differential attrition presumably occurred because participants received many more text messages during the four-week intervention in the two *information provision* conditions than in the *no-treatment control* condition (34.5 messages in the *information provision* conditions, on average vs. 16 in the *no-treatment control* condition, on average; see Figure 6 for a visualization of attrition rates over time and as the number of texts accumulated across conditions). Participants in the *no-treatment control* condition may have experienced less text messaging fatigue, which could explain why they were both less likely to opt out of the experiment and more likely to respond to our text messages (Fricke, Kalogrides, & Loeb, 2018).

Importantly, this differential attrition does not significantly impact our ability to draw inferences from the results of Study 2. Because we conducted all analyses with respect to intent-to-treat effects, they are not vulnerable to selection effects: all participants, regardless of whether they remained in the intervention, are included in our statistical tests. Moreover, we used the most pessimistic possible assumptions for participants who opted out, assuming they would not adhere to public health recommendations, would not seek more information, and would not share more information. In other words, we operationalized opting out as a signal of being completely unaffected by the intervention. These pessimistic assumptions ultimately yield two conclusions: first, sharing public health information via text message may not be the most effective strategy if keeping people subscribed to text messages is the ultimate goal, given that the *no-treatment control* condition seems to outperform the two *information provision* conditions by retaining more subscribers (and therefore generating more information-seeking and self-reported behavior change). Second, relative to sharing information in direct statement format, sharing it in a Q&A format is more likely to increase engagement and adherence with recommended behaviors. This second comparison is our focal comparison of interest, both theoretically and practically. Theoretically, the tightly matched *Q&A* and *direct statement* conditions allow us to make inferences about the benefits of inducing curiosity when attempting to trigger engagement with information.

Practically, sharing information is often an imperative for governments and managers, making the holdout control an irrelevant comparison condition. Further, managers and policymakers most commonly share information in direct statement format, which makes this condition the right comparison condition for assessing whether Q&A-style communication outperforms current best practices.

4. Study 3: Implementation Experiment on Facebook

In Study 3, we aim to further probe the practical value of our results from Studies 1 and 2 by running a public health ad campaign encouraging vaccination using Facebook's A/B testing software. Specifically, in an ad campaign with fixed daily spend set to optimize cost-per-click, we tested the relative cost-effectiveness of two ads: one that posed a question about COVID-19 vaccines and required users to click a link to learn the answer (the *Q&A* ad) and another that directly shared the corresponding fact about vaccines and urged users to click the link to learn more (the *direct statement* ad). Note that cost-effectiveness is measured in terms of unique link clicks, not in terms of the ultimate behavior of interest (e.g., getting a vaccine), which is unobservable to us.

4.1 Methods

We designed a Facebook ad campaign that reminded users about the social benefit of being vaccinated to promote engagement with accurate information about COVID-19 vaccines. Our campaign ran for 27 days (from June 11, 2021 until July 8, 2021) and reached an estimated 947,818 Facebook users.[21]

[21] While this study was originally preregistered on AsPredicted.org (https://aspredicted.org/h69z-jnd7.pdf), we had to deviate from our planned preregistration when we learned that randomization was not occurring at the level we originally expected. When planning our data analysis, we believed that Facebook would be randomly assigning users to view different ads. Instead, Facebook randomly divides users from the target audience into two groups before launching the ad campaign, and these become the potential audiences for each version of the ad. Which potential audience members actually see the ad (and how many times) is determined by an internal bidding strategy that considers the total amount of money the advertiser is willing to spend per day and the effectiveness of the ad amongst different sub-segments of the potential audience population. Thus, while Facebook initially divides the target groups into balanced, random, non-overlapping audiences, they then use optimization strategies to show each ad primarily to sub-segments of the respective overall audience group that are most promising. As a result, there are imbalances across groups assigned to the two experimental conditions, including imbalances in the frequency with which users saw each ad and the number of users who saw each ad (see Braun et al., 2024 or Braun & Schwartz, 2024 for a discussion of Facebook's A/B testing strategy and its implications for causal inference). Thus, we cannot compare the dependent variable we originally planned to compare (click-through rate at the individual level) across conditions. Instead, we can and do compare these advertisements on cost effectiveness; in other words, we compare

We used Facebook's internal A/B testing system to compare the performance of a *Q&A* ad and a *direct statement* ad (both described below). The campaign was set up to optimize for link clicks, with cost-per-click as the key outcome metric. The A/B testing system randomizes users from the target population into two non-overlapping, balanced groups, each associated with one of the two ads; the system is designed to assess which ad will generate greater clicks per dollar spent among the target population. To do so, the A/B testing system holds fixed the daily dollars spent (budget) across the two groups. For each group, Facebook's algorithms automatically bid in Facebook's advertising auctions with the goal of obtaining as many clicks as possible within the ad's daily budget. We measured the number of unique link clicks per day each ad generated, noting that since the dollars spent per day is equal across the ads, a treatment effect on clicks per day can be translated into a treatment effect on clicks per dollar simply by dividing by the dollars spent per day.

The ads were launched to adult U.S. users only (our target population). Users were randomly divided into two non-overlapping, balanced audiences: a *Q&A* ad audience and a *direct statement* ad audience. Facebook's internal bidding system determined which subset of each audience population was exposed to each ad and how many times each user saw the ad assigned to them. Overall, 489,665 unique users saw the *Q&A* ad and 458,153 saw the *direct statement* ad, where the number of users should be

---

the unique clicks each advertisement generated conditional on daily spend (About A/B Testing, Facebook Business Help Center, n.d.). We can assess the effect of ad type on total unique clicks because the daily spend is held constant across ads, the initial potential audience is randomly assigned, and both ads are subject to the same optimization protocol (i.e., Facebook's algorithm will try to ensure that each ad gets as many clicks as possible given its daily budget). We analyze unique clicks as this was the metric our ad campaign sought to optimize, and we control for the daily spend on each ad in our regression, meaning that comparisons of unique clicks are essentially equivalent to comparing clicks per dollar spent. Using this DV, our results are high on both internal validity and ecological validity: real advertisers trying to use this strategy to improve engagement with health information will encounter similar platform optimization when posting ads. Moreover, the results complement the findings from Studies 1 and 2, and this convergent evidence reduces the likelihood that the effects we find are spurious.

To achieve identification of average individual level effects, social media researchers may wish to consider alternative designs. Advertisements placed on Facebook are subject to internal bidding strategies which are black boxes to researchers; when this bidding is downstream of randomization, it is not possible for researchers to ensure balance in delivery of advertisements across conditions. Researchers may instead wish to use designs where bidding is upstream of randomization, for example in messaging experiments where users are recruited through advertisements, engaged in messaging conversations, and then assigned treatment. However, interventions that are most relevant to public health campaigns may not be testable in such contexts; instead, they require designs like Study 3's, where there is variation at the advertisement level.

considered an outcome that is influenced by Facebook's estimates of the quality of the ad and of the most efficient way to allocate our fixed budget to attract clicks.

The *Q&A* ad prompted users to consider a question and required them to click the link in the ad to learn the answer: "Can you see your vaccinated friends mask-free if you're fully vaccinated?" The *direct statement* ad, on the other hand, directly informed users of the answer: "The CDC says it's safe to socialize mask-free with your friends if you're fully vaccinated, but only if they're fully vaccinated, too." (see Figure 7). Across both ads, users were prompted to "click here to learn more about vaccines," and the link redirected them to a CDC webpage about vaccines (Center for Disease Control and Prevention, 2021). Given this directive, we interpret unique clicks as information-seeking behavior.

4.2 Results

We relied on the following log-linear regression with robust standard errors to identify the effects of the *Q&A* ad relative to the *direct statement* ad:

$$ln\,(Y_{a,d}) = \beta Q_a + \gamma ln\,(S_{a,d}) + \alpha D_d + \varepsilon_{a,d}$$

$Y_{a,d}$ is the number of unique link clicks generated by ad *a* on day *d* of the ad campaign, Q_a is an indicator for whether ad *a* is the *Q&A* ad, $S_{a,d}$ is the amount spent on ad *a* on day *d*, and D_d is an indicator for the day *d* of the ad campaign. Note that this analysis was conducted at the day level rather than the individual level because randomization did not occur at the individual level (and, as expected, there are imbalances across conditions due to Facebook's randomization algorithm, see Footnote 21 and Supplement Table 19). We aggregate results at the daily level because the ad campaign had a daily budget, allowing each day to be considered an observation of the results of bidding a fixed amount of money on the *Q&A* ad relative to the *direct statement* ad.

While our ad campaign ran for 27 days, there was substantial variation over time in our outcome (see Supplement Figure 2). During the early stages of the A/B test while data was still being gathered, both ads' performance fluctuated as Facebook learned how to target the ads. Moreover, there was substantial variability in spend, $S_{a,d}$, within and across days at the beginning of the ad campaign (due to the burn-in period) and at the end of the ad campaign (as our campaign budget ran out). As a result, we present our results considering two different time periods: (1) the entire 27-day ad campaign and (2) the 18 days from June 18, 2021 to July 5, 2021 when both daily spend and our outcome were stable.

Consistent with Studies 1 and 2, and as shown in Table 6, Model 1, the *Q&A* ad inspired a 9.1% increase in unique link clicks relative to the *direct statement* ad (b = 0.09, SE = 0.01, $p < 0.001$; daily average unique link clicks: *Q&A* ad = 3166.0, *direct statement* ad = 2891.5) when considering our full dataset. If we restrict our attention to after the burn-in period and before our budget started to run out (as shown in Table 6, Model 2), the *Q&A* ad generated an 11.3% increase in unique link clicks relative to the *direct statement* ad (b = 0.11, SE = 0.01, $p < 0.001$; daily average unique link clicks: *Q&A* ad = 3755.2, *direct statement* ad = 3369.3).[22] In other words, a practitioner with a fixed budget interested in encouraging people to click through an ad to access more public health information could expect to engage 9.1%-11.3% more unique social media users with a Q&A-style ad than with a direct statement-style ad.

4.3 Discussion

Study 3 provides evidence that piquing curiosity with a question is a useful communication strategy for social media-based ad campaigns. This study demonstrates that delivering information in Q&A format can be useful in contexts beyond text messaging, and in particular, is more cost effective when the goal is to induce engagement with the information (in this case, by clicking on a link). On the other hand, the *direct statement* ad may have shifted behavior of people who did not click on the link since the ad contained a fact and not just a question; hence the ultimate impact of the two ads on people's

[22] We present the results of robustness checks in Supplement Tables 20 and 21 in which we (1) analyze the data using a Poisson regression (estimated treatment effect overall = 9.7%, SE = 0.01; estimated treatment effect after the burn-in period = 11.4%, SE = 0.01) and (2) adjust for serial correlation (estimated treatment effect overall = 6.8%, SE = 0.02, estimated treatment effect after the burn-in period = 11%, SE = 0.01). Our results are robust to further filtering to focus only on the 12 days between June 24, 2021 and July 5, 2021 (estimated treatment effect = 11%, SE =0.02). To explore further robustness checks, interested readers can use the code posted on OSF (see: https://osf.io/5ab26/).

subsequent behaviors is ambiguous. Future work measuring the impact of such ads on behavior would be valuable.

5. General Discussion

Across two large, preregistered field experiments conducted in Ghana and Michigan and an implementation experiment on Facebook, we show that delivering information in a Q&A format increases engagement with and adherence to recommended health behaviors. Ghanaian citizens who were texted COVID-19 facts in a Q&A format rather than as direct statements were 43% more likely to opt in to further information seeking opportunities ($p < 0.001$; lift in information-seeking: participants in our *Q&A* condition took 0.02 out of 5 more information-seeking opportunities than those in our *direct statement* condition). Exploratory analyses suggest this was primarily driven by opportunities to seek information about specific topics (e.g., mask wearing) rather than opportunities to access resources with general information about COVID-19. Similarly, Michigan residents who received information in Q&A format rather than as direct statements were 19% more likely to seek further topic-specific information ($p < 0.001$; lift in specific information-seeking: participants in our *Q&A* condition took 0.03 more information-seeking opportunities out of 3 than those in our *direct statement* condition), but they were 12% less likely to seek access to general information about COVID-19 ($p < 0.001$; drop in general information-seeking: participants in our *Q&A* condition took 0.02 fewer general information-seeking opportunities out of 2 than participants in our *direct statement* condition). Michigan residents who received information in Q&A format were also 4.4% more likely ($p = 0.002$) to report adhering fully to public health guidelines than those who received information in direct statement format (lift in reported adherence to recommended health behaviors: participants in our Q&A condition reported adhering to 0.05 more behaviors out of 4 than participants in our direct statement condition), suggesting that piquing curiosity not only increases interest in future learning opportunities but also increases people's willingness to apply information to their own lives. In a follow-up implementation study, we found that Q&A-style messaging increased the cost-effectiveness of public health ads on social media: an ad containing a question about COVID-19 vaccines rather than a fact about vaccines inspired a 9.1%-11% increase in the number of unique

Facebook users who clicked a link to learn more. Together, these findings suggest that across several communication channels, populations, and cultures, Q&A-format information delivery can be a fruitful way to encourage people to engage with public health information.

While our effect sizes in Studies 1 and 2 are small (1 to 1.5 percentage-point increases in information seeking and self-reported behavior change), it is promising that a low-cost, scalable change (adding a question to a text message sequence, which typically costs $0.01 in the U.S. and $0.11 in Ghana) can reliably increase engagement with critical public health information (Benartzi et al., 2017; List, 2024). Relative to the status quo—sending critical public health information via text in direct statement format—the total extra cost per-person of Q&A-style messaging in our studies was $0.88 in Ghana and $0.08 in the U.S. The effect sizes we document suggest that every marginal $1,000 spent by governments to support Q&A style messaging (instead of direct statement messaging) would lead to 24 extra citizens engaging with public health information in Ghana, but 300 in Michigan. Spending an extra $1,000 on Q&A texts in Michigan (instead of direct statement texts) would also lead to 650 more citizens confirming that they engaged in a recommended health behavior at least once; for example, for every extra $1,000 spent in Michigan on Q&A messages (as opposed to direct statement texts), an extra 257 residents reported social distancing. Overall, Q&A-style interventions are more cost-effective in the U.S. context than in Ghana due largely to the higher cost of texting in Ghana. However, Study 3 suggests that insights from this research can be useful not only for improving informational text messaging campaigns, but also for enhancing advertising (e.g., on social media, where switching to a Q&A frame from a direct statement frame significantly *reduces* costs). To maximize the impact of each dollar spent, policymakers and managers should keep an eye on costs and platform penetration to decide how and whether to implement Q&A-style messaging in their context.

To benchmark the effect sizes we document relative to those achieved by other RCTs testing the benefits of policy interventions relying on "nudges" (often reminders or reframed communications), we refer to DellaVigna and Linos's (2022) review of 126 RCTs run by Nudge Units affiliated with the U.S. government. They document an average treatment effect of 1.49 percentage-points across all 126 RCTs, meaning that a target behavior—say, filing taxes on time—increases by 1.49 percentage points, on average, in response to a nudge. Of particular relevance for us is the reported average treatment effect across the 24 RCTs focused on changing public health behavior: here the average effect is a 0.49 percentage-point improvement in the targeted behavior. The modest nature of this effect size illustrates how difficult public health behaviors can be to move and suggests that the effect sizes we detect (of 1-1.5 percentage-point changes in targeted behaviors) are roughly twice as large as the typical impact of a nudge targeting public health.

Our field experiment with Michigan residents also highlights an important limitation of text messaging to share valued information with a key constituency: repeated messaging can lead people to disengage fully. People who received numerous text messages with public health information—whether in Q&A format or via direct statements—were 69% more likely ($p < 0.001$) to fully opt out of receiving future text messages compared to a control group that received no information. This is consistent with past research showing that people are more likely to opt out of informational text-messaging interventions when they receive (1) a higher quantity of messages and (2) more complex messages, both of which can increase text messaging fatigue (Fricke et al., 2018; Steiner et al., 2021). However, our intent-to-treat analyses result in internally valid estimates that combine potentially divergent effects of Q&A-style text communications on behavior.

Our findings add nuance to a substantial prior literature suggesting that friction-reduction is optimal for encouraging uptake of important information (e.g., Handel & Schwartzstein, 2018; Bergman et al., 2020; Fishbane et al., 2020; Lasky-Fink et al., 2021; Shah & Oppenheimer, 2008). This literature suggests that the most effective way to combat attention scarcity is to maximize ease of accessing and processing whatever information is being delivered, thereby reducing the likelihood that important facts are overlooked. However, our work suggests that adding *some* friction may be beneficial when doing so also stimulates curiosity or signals that the information is of value. Although we do not explicitly test this mechanism in our work, we theorize that because curiosity drives attention and increases enjoyment in the learning process (Ruan et al., 2018; Kang et al., 2009; Buyalskaya & Camerer, 2020), it increases the likelihood that people will engage with, recall, and utilize new information. Future work should explore whether Q&A-style communication could be rendered even more effective via gamification (for example, by adding a point system or symbolic rewards to the Q&A messages), which might further capitalize on the fun that curiosity injects into the learning process (Mollick & Rothbard, 2014; Patel et al., 2019).

However, our results suggest that the effects of curiosity on engagement may be narrow, which speaks to competing predictions in the (primarily lab-based) literature on curiosity. Given that curiosity boosts intrinsic motivation to learn, it might be reasonable to predict that it would broadly increase interest in acquiring new information about a general topic (Ruan et al., 2018). However, sharing information in a Q&A format in our field experiments in Ghana and in Michigan only increased people's interest in learning more about specific, narrow topics related to the facts we had just shared, and not in accessing resources with more general information. This is consistent with prior theorizing suggesting that when someone learns more about a specific topic, the information they *lack* is also brought into focus, further increasing their desire to acquire information about that topic; meanwhile, because information gaps about other topics (even related topics) are not salient, people are unlikely to seek more information about them (Loewenstein, 1994). Although they are only suggestive, our results seem to fall in line with this prediction.

Specifically, our results lead us to speculate that curiosity may act more like a flashlight than a lightbulb, shining a beam of attention on a narrow radius of relevant information rather than lighting up a diffuse, generalized desire to learn. For example, imagine that a friend asks you which animals can draw self-portraits. Once you learn that (one) answer to this question is "elephants," a lightbulb model suggests that you might be excited to learn more about animals in general, leading you to seek out more fun facts about the animal kingdom. A flashlight model, on the other hand, would suggest that you will want to know more about elephants *in particular*, seeking out information about their artistic, social, and communication habits. While our studies do not explicitly test this question, the flashlight model seems to be better supported by our results: asking a question about mask-wearing motivates people to learn more about masks specifically, but not about COVID-19 in general. We note, however, that this analysis is purely speculative: our experiments were not explicitly designed to test this question, so we cannot rule out alternative explanations for this difference. For example, specific information-seeking opportunities always immediately followed a fact shared in either Q&A format or as a direct statement (see Figures 1 and 2, Supplement Tables 19 and 20), whereas general information seeking opportunities were sent on days when participants had not learned a new fact. It may be that the effects of stimulating curiosity are not topically narrow but temporally narrow, fading quickly over time. Alternatively, participants in the Q&A condition may have simply been exhibiting the effects of momentum—some of them had just interacted with our texts to receive an answer to a question, making it more likely that they would interact again when another information-seeking opportunity presented itself immediately afterwards. Future work should explore whether piquing curiosity increases information seeking about broadly vs. narrowly related topics, and whether the effects of stimulating curiosity are ephemeral or lasting.

Relatedly, future work is needed to test whether Q&A-style communication increases sustained attention to new information. While our studies confirm that "teasing" people with a question before sharing facts can increase momentary attention (i.e., inducing people to click on a link or text "1" to learn more), we cannot confirm whether there were also long-term benefits for willingness to learn new information (Berger et al., 2023). As an initial assessment, we might assume that participants who exhibited sustained attention to the facts we shared would be more likely to display subsequent adherence to recommended health behaviors (because they would be more likely to recall and apply the information they learned). This analysis yields mixed conclusions. We do not find evidence that Q&A-style communication boosted willingness to wear masks in Michigan, even though participants learned about masks during the first week of the intervention. However, we do find increased adherence with public health recommendations about hand-washing and social distancing during the last two weeks of the intervention, each of which were also preceded by relevant facts. We encourage future work to further explore the impact of Q&A messaging on sustained attention.

While our theorizing focuses on curiosity as a likely primary driver of the benefits of Q&A-style communication, our results may also be driven by an alternative (or complementary) mechanism. In our experiments, the information-delivery method that we designed to pique people's curiosity also required more effort. Participants who received information in direct statements were simply presented with facts, while participants who received information in a Q&A format had to answer a text, click on a link, or wait 24 hours to access the same fact. Though prior work cautions that increasing the effort required to obtain information is likely to decrease uptake and engagement, it may be that those who do exert effort are subsequently more engaged (Bergman et al., 2020). People are more likely to recall information they have been asked to self-generate (e.g., by guessing an answer to a question) than information they simply read, in part because they have exerted more effort to acquire the information (Jurica & Shimamura, 1999; for a meta-analysis, see Bertsch et al., 2007). Effort may do more than improve recall: taking an active role in the information-acquisition process may lead people to infer that the information they are acquiring is more valuable, increasing their subsequent interest in learning more (Inzlicht et al., 2018; Olivola & Shafir, 2013; Norton et al., 2012; Kim & Labroo, 2011). Future work should explore whether requiring participants' effort is a necessary ingredient for successful Q&A format information delivery. For example, it might be worth testing whether effort-free Q&A information delivery (e.g., an FAQ) is as effective at promoting interest in new information to disentangle whether curiosity or effort drives the benefits of Q&A-style communication.

Several limitations of this work are worth noting. First, in Ghana, response rates to our texts were quite low (at or below 1.5%), likely due to study participants' (misplaced) concerns about paying text messaging fees (see "Additional Implementation Details" in the Supplement). This prevented us from gathering usable insights about adherence to recommended health behaviors across conditions, as we cannot infer whether the ~99% of participants who did not respond to questions adhered to public health guidelines or not.

Second, because participants in both Ghana and Michigan had to opt into our text-messaging intervention, our experiments are internally valid but may have more limited external validity to populations without an expressed interest in (or concerns about) COVID-19. As a result, we may be measuring the effectiveness of Q&A formatted messaging in a population of people with a high baseline level of interest in the information being shared rather than in the overall population. Our insights should still be useful to many practitioners, however. Requiring people to opt in to receive text messages is common for legal reasons, so investigations of the impact of texting on public health behavior frequently rely on opt-in consent (e.g., Milkman et al., 2021; Dai et al., 2021; Milkman et al., 2022; Lewey et al., 2022).

Third, in Michigan, a large proportion of participants opted out of our intervention, leading to significant differential attrition across conditions. Because all analyses were conducted intent-to-treat, and because there was more attrition in our Q&A condition than our direct statement condition, we may be underestimating the benefits of sharing information in Q&A format. In fact, the benefits of Q&A-style communication we document seem driven by particularly high rates of engagement among participants who did not choose to opt out. This suggests that sharing information in Q&A format (at least via text message) may lead to divergent responses, with some people disengaging completely and others exhibiting increased engagement. Heterogeneous responses to Q&A-style information delivery are worth investigating in future work.

Fourth, we sent more text messages in the Q&A condition than in the direct statement condition in Studies 1 and 2. These extra text messages could have served as a boon, increasing people's engagement with our intervention (because they received more frequent communications) or reinforcing the information we shared (because people first inferred the answer, then had it confirmed). However, given that attrition rates in Study 2 increased as a function of the number of texts participants received (see Figure 6), we think it is more likely that the extra messages alienated participants. Moreover, differences in number and frequency of messages cannot explain the benefits of Q&A-style communication in Study 3. However, we acknowledge that the efficacy of the Q&A condition relative to the direct statement condition in Studies 1 and 2 could partially be driven by this difference in text message frequency and disentangling this possibility would be a valuable direction for future research.

Fifth, our studies suffer from measuring primarily intermediate outcomes (e.g., information-seeking behaviors) or noisy behavioral measures (e.g., self-reported behaviors) rather than realized behavior change or health outcomes (e.g., vaccine uptake). Though past work suggests that increased engagement is linked to behavior change (Schwarzer & Satow, 2012; Tan et al., 2012; Sheeran et al., 2017), future work should confirm the impact of Q&A-style communications on downstream health decisions. For example, future studies could re-run our implementation experiment with an outcome that is more proximal to the final behavior of interest (e.g., click-through to a page with vaccine sign-ups, or completion of such a page).

Our findings suggest that managers, policymakers, and advertisers seeking to communicate critical information should craft messages that "tease" readers with a question prior to sharing information, boosting recipients' curiosity to maximize engagement. Human resource managers who need employees to understand key changes to their benefits elections, for example, might try sharing an email with a question about employee benefits in the subject line rather than a statement. A scientist communicating the result of research on climate change to policymakers may be better off sharing an FAQ document rather than a bulleted list of key facts. And a manager hoping to inform direct reports about a new learning and development module might be better off sending questions (and, eventually, answers) about the new module via Slack than simply writing up a paragraph to describe it. In the current information-rich environment, might questions spur people to direct their limited attention to information that can improve their lives? Our experiments suggest the answer is "yes."

References

About A/B Testing. (n.d.). Facebook Business Help Center. https://www.facebook.com/business/help/1738164643098669?id=445653312788501

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., ... & Galing, S. (2017). Should governments invest more in nudging?. *Psychological Science*, *28*(8), 1041-1055.

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. *Journal of Marketing*, *87*(5), 793–809.

Bergman, P., Lasky-Fink, J., & Rogers, T. (2020). Simplification and defaults affect adoption and impact of technology, but decision makers do not realize it. Organizational Behavior and Human Decision Processes, 158, 66-79.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, *35*(2), 201-210.

Braun, M., De Langhe, B., Puntoni, S., & Schwartz, E. M. (2024). Leveraging digital advertising platforms for consumer research. *Journal of Consumer Research*, *51*(1), 119–128.

Braun, M., & Schwartz, E. M. (2024). Express: Where a-b testing goes wrong: how divergent delivery affects what online experiments cannot (And can) tell you about how customers respond to advertising. *Journal of Marketing*, 00222429241275886.

Buyalskaya, A., & Camerer, C. F. (2020). The neuroeconomics of epistemic curiosity. *Current Opinion in Behavioral Sciences*, *35*, 141-149.

Centers for Disease Control and Prevention. (2021). *Key Things to Know About COVID-19 Vaccines*. Centers for Disease Control and Prevention. Retrieved June 1, 2021, from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/keythingstoknow.html

Dai, W., & Luca, M. (2020). Digitizing disclosure: The case of restaurant hygiene scores. *American Economic Journal: Microeconomics*, *12*(2), 41-59.

Dai, H., Saccardo, S., Han, M. A., Roh, L., Raja, N., Vangala, S., ... & Croymans, D. M. (2021). Behavioural nudges increase COVID-19 vaccinations. *Nature*, *597*(7876), 404-409.

DellaVigna, S., & Linos, E. (2022). Rcts to scale: Comprehensive evidence from two nudge units. *Econometrica*, *90*(1), 81–116.

Department of Global Communications. (2020). *'Verified' initiative aims to flood digital space with facts amid COVID-19 crisis*. United Nations. https://www.un.org/en/coronavirus/%E2%80%98verified%E2%80%99-initiative-aims-flood-digital-space-facts-amid-covid-19-crisis

Digital Health and Innovation Team. (2020). *COVID-19 SMS message library*. World Health Organization. https://www.who.int/publications/i/item/covid-19-message-library

Fishbane, A., Ouss, A., & Shah, A. K. (2020). Behavioral nudges reduce failure to appear for court. *Science*, *370*(6517), eabb6591.

Fricke, H., Kalogrides, D., & Loeb, S. (2018). It's too annoying: Who drops out of educational text messaging programs and why. *Economics Letters*, *173*, 39-43.

Ghana Statistical Service. (2013). *2010 Population & Housing Census: Demographic, Social, Economic & Housing Characteristics*. Ghana Statistical Service.

Handel, B., & Schwartzstein, J. (2018). Frictions or mental gaps: What's behind the information we (don't) use and when do we care?. *Journal of Economic Perspectives*, *32*(1), 155-78.

Horn, S., Litovsky, Y., & Loewenstein, G. (2024). Using curiosity to counter health information avoidance. *Social Science & Medicine*, *340*, 116383.

Hsee, C. K., & Ruan, B. (2016). The Pandora effect: The power and peril of curiosity. *Psychological Science*, *27*(5), 659-666.

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly

and valued. *Trends in Cognitive Sciences*, *22*(4), 337-349.
Jin, G. Z., Luca, M., & Martin, D. (2021). Is no news (perceived as) bad news? An experimental investigation of information disclosure. *American Economic Journal: Microeconomics*, *13*(2), 141-73.
Jin, G. Z., Luca, M., & Martin, D. (2022). Complex disclosure. *Management Science*, *68*(5), 3236-3261.
John, L. K., Blunden, H., Milkman, K. L., Foschini, L., & Tuckfield, B. (2022). The limits of inconspicuous incentives. *Organizational Behavior and Human Decision Processes*, *172*, 104180.
Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition*, *27*(4), 648-656.
Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T. Y., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963-973.
Kim, S., & Labroo, A. A. (2011). From inherent value to incentive value: When and why pointless effort enhances consumer preference. *Journal of Consumer Research*, *38*(4), 712-742.
Kostyk, A., Leonhardt, J. M., & Niculescu, M. (2021). Processing fluency scale development for consumer research. *International Journal of Market Research*, *63*(3), 353–367.
Kupor, D. M., & Tormala, Z. L. (2015). Persuasion, interrupted: The effect of momentary interruptions on message processing and persuasion. *Journal of Consumer Research*, *42*(2), 300-315.
Lasky-Fink, J., Robinson, C. D., Chang, H. N. L., & Rogers, T. (2021). Using behavioral insights to improve school administrative communications: The case of truancy notifications. *Educational researcher*, *50*(7), 442-450.
Lewey, J., Murphy, S., Zhang, D., Putt, M. E., Elovitz, M. A., Riis, V., ... & Levine, L. D. (2022). Effectiveness of a text-based gamification intervention to improve physical activity among postpartum individuals with hypertensive disorders of pregnancy: a randomized clinical trial. *JAMA Cardiology*, *7*(6), 591-599.
Loewenstein, G., Prelec, D., & Shatto, C. (1998). Hot/cold intrapersonal empathy gaps and the under-prediction of curiosity. *Unpublished manuscript, Carnegie-Mellon University, Pittsburgh, PA*.
Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75.
Loewenstein, G., & Wojtowicz, Z. (2023). *The economics of attention* [SSRN Scholarly Paper]. Rochester, NY.
Milkman, K. L., Gandhi, L., Patel, M. S., Graci, H. N., Gromet, D. M., Ho, H., ... & Duckworth, A. L. (2022). A 680,000-person megastudy of nudges to encourage vaccination in pharmacies. *Proceedings of the National Academy of Sciences*, *119*(6), e2115126119.
Milkman, K. L., Patel, M. S., Gandhi, L., Graci, H. N., Gromet, D. M., Ho, H., ... & Duckworth, A. L. (2021). A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proceedings of the National Academy of Sciences*, *118*(20), e2101165118.
Mollick, E. R., & Rothbard, N. (2013). Mandatory fun: Gamification and the impact of games at work. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2277103
Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, *22*(3), 453–460.
Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making*, *26*(1), 91-105.
Patel, M. S., Small, D. S., Harrison, J. D., Fortunato, M. P., Oon, A. L., Rareshide, C. A. L., … Hilbert, V. (2019). Effectiveness of behaviorally designed gamification interventions with social

incentives for increasing physical activity among overweight and obese adults across the united states: The step up randomized clinical trial. *JAMA Internal Medicine*, *179*(12), 1624.
Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590-595.
Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39-50.
Polman, E., Ruttan, R. L., & Peck, J. (2022). Using curiosity to incentivize the choice of "should" options. *Organizational Behavior and Human Decision Processes*, *173*, 104192.
Ruan, B., Hsee, C. K., & Lu, Z. Y. (2018). The teasing effect: An underappreciated benefit of creating and resolving an uncertainty. *Journal of Marketing Research*, *55*(4), 556-570.
Schwarzer, R., & Satow, L. (2012). Online intervention engagement predicts smoking cessation. *Preventive Medicine*, *55*(3), 233-236.
Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological bulletin*, *134*(2), 207.
Sheeran, P., Klein, W. M., & Rothman, A. J. (2017). Health behavior change: Moving from observation to intervention. *Annual Review of Psychology*, *68*, 573-600.
Shen, L., Fishbach, A., & Hsee, C. K. (2015). The motivating-uncertainty effect: Uncertainty increases resource investment in the process of reward pursuit. *Journal of Consumer Research*, *41*(5), 1301-1315.
Steiner, J. F., Zeng, C., Comer, A. C., Barrow, J. C., Langer, J. N., Steffen, D. A., & Steiner, C. A. (2021). Factors associated with opting out of automated text and telephone messages among adult members of an integrated health care system. *JAMA Network Open*, *4*(3), e213479-e213479.
Tan, A. S., Mello, S., & Hornik, R. C. (2012). A longitudinal study on engagement with dieting information as a predictor of dieting behavior among adults diagnosed with cancer. *Patient Education and Counseling*, *88*(2), 305-310.
Wojtowicz, Z., & Loewenstein, G. (2020). Curiosity and the economics of attention. *Current Opinion in Behavioral Sciences*, *35*, 135-140.
Woolley, K., & Fishbach, A. (2016). For the fun of it: Harnessing immediate rewards to increase persistence in long-term goals. *Journal of Consumer Research*, *42*(6), 952-966.
Yang, T.-M., & Maxwell, T. A. (2011). Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors. *Government Information Quarterly*, *28*(2), 164–175.

Table 1. Design features that varied across experiments.

	Site	Message Crafting	Sample size	Recruitment Method	Time between recruitment and intervention start date	Duration of intervention	Dependent variables measured	Conditions included
Study 1	Ghana	Crafted messages in collaboration with a focus group in Ghana (see Supplement for more details)	11,585	Phone survey and Interactive Voice Response (IVR) phone calls placed with random digit dialing	Varied across participants (start day was the same for all participants, lag time depended on when they opted in)	7.5 weeks (53 days between December 10, 2020 and February 1, 2021)	(1) Information seeking a. Opportunities to seek access to information about specific topics, e.g., mask wearing b. Opportunities to seek access to a phone number you can call for free, general information about COVID-19 (2) Information sharing	(1) Q&A condition (2) Direct statement condition
Study 2	Michigan	Crafted messages with the Michigan Department of Health and Human Services (see Supplement for more details)	29,810	Recruitment messages send to numbers provided by the Michigan Department of Health and Human Services	1 day (start day was staggered across participants based on when they opted in)	4 weeks (start day ranged from November 13, 2020 and December 31, 2020)	(1) Self-reported actions taken to prevent the spread of COVID-19 (2) Information seeking a. Opportunities to seek access to information about specific topics, e.g., mask wearing b. Opportunities to seek access to resources with general information about COVID-19 (e.g., myth-busting websites) (3) Information sharing	(1) Q&A condition (2) Direct statement condition (3) No-treatment control condition

Note. This table describes implementation details across Studies 1 and 2. Note that all messages sent in Studies 1 and 2 were vetted by a group of doctors and public health officials for accuracy prior to our experiments. Because Study 3 was an implementation experiment rather than a traditional RCT, it was not included in this table.

Table 2. Regression-Estimated Effects of Q&A Treatment on Composite Information Seeking and Information Sharing in Study 1.

	Model 1 *Outcome: Information Seeking Composite*			**Model 2** *Outcome: Topic-Specific Information Seeking Composite*			**Model 3** *Outcome: General Information Seeking Composite*			**Model 4** *Outcome: Number of phone numbers shared*		
	b	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**
Q&A Condition	0.023	[0.011, 0.035]	<.001	0.028	[0.020, 0.035]	<.001	-0.004	[-0.012, 0.003]	0.259	0.001	[-0.000, 0.003]	.104
Days elapsed between recruitment and start of intervention	-0.000	[-0.000, -0.000]	<.001	-0.000	[-0.000, - 0.000]	<.001	-0.000	[-0.000, - 0.000]	<.001	-0.000	[-0.000, -0.000]	.011
Observations		11585			11585			11585			11585	
Adjusted R^2		0.004			0.007			0.002			0.001	

Note. This table reports the results of four ordinary least squares (OLS) regression models predicting information seeking and information sharing in Study 1. The first regression model predicts a given Study 1 participant's information seeking composite score, generated by summing the number of times (out of 5) the participant responded affirmatively to an information seeking opportunity (Model 1, preregistered). The second regression model predicts a given Study 1 participant's topic-specific information seeking composite score, generated by summing the number of times (out of 3) the participant responded affirmatively to an offer of more information about a specific topic related to COVID-19 (Model 2, exploratory). The third regression model predicts a given Study 1 participant's general information seeking composite score, generated by summing the number of times (out of 2) the participant responded affirmatively to an opportunity to access a resource with general information about COVID-19 (Model 3, exploratory). The fourth regression model predicts the number of phone numbers of people they would like to keep informed about COVID-19 a given participant in Study 1 shared (Model 4, preregistered). All models show the main effect of assignment to the Q&A condition. The models also include a continuous control for the number of days prior to the intervention's start date that the participant was recruited.

Table 3. Regression-Estimated Effects of Q&A Treatment Relative to Direct Statements on Composites Measuring Self-Reported Adherence to Recommended Health Behavior in Study 2.

	Model 1 *Outcome: Behavior Composite*			**Model 2** *Outcome: Behavior Adherence Composite, Binary (Exploratory)*		
	b	**95% CI**	**p**	**b**	**95% CI**	**p**
Q&A Condition	0.035	[0.016, 0.053]	<.001	0.053	[0.019, 0.086]	.002
Days Between Intervention Launch and Participant Opt-in	-0.002	[-0.003, -0.001]	<.001	-0.004	[-0.005, -0.003]	<.001
Observations		29810			29810	
Adjusted R^2		0.005			0.005	

Note. This table reports the results of two ordinary least squares (OLS) regression models predicting self-reported adherence to recommended health behaviors among Study 2 participants. Comparisons between the *Q&A* and *direct statement* conditions are derived from Wald tests, as preregistered. The first regression model predicts a given Study 2 participant's preregistered behavior change composite score, generated by z-scoring and averaging responses to each of the four self-reported behavior measures, replacing missing values with the "worst" possible response (e.g., maximal non-compliance) (Model 1, preregistered). The second regression model predicts a given Study 2 participant's behavior adherence composite score, generated by transforming non-binary self-reported behaviors into a binary DV (complied vs. did not comply, with missing values coded as non-compliance), and summing the number of times, out of four, that participants complied (Model 2, exploratory). All models show the main effect of assignment to the *Q&A* condition. The models also include a continuous control for the day the participant was recruited and an indicator for assignment to the *no-treatment control* condition (not shown here; see Supplement Table 11 for results comparing the *no-treatment control* condition to the *information provision* condition).

Table 4. Regression-Estimated Effects of Q&A Treatment Relative to Direct Statements on Self-Reported Behaviors in Study 2.

	Model 1 *Outcome: Did the participant report always wearing a masking when leaving home? (Day 5)*			**Model 2** *Outcome: How many of the past 7 days did the participant report leaving home for non-essential reasons? (Day 12)*			**Model 3** *Outcome: Did participant report washing their hands thoroughly the last time they returned from an outing? (Day 20)*			**Model 4** *Outcome: Did participant report touching, hugging, or shaking hands with anyone outside household? (Day 28)*		
	b	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**
Q&A Condition	-0.015	[-0.028, -0.003]	.018	0.022	[-0.061, 0.105]	.605	0.049	[0.038, 0.059]	<.001	0.021	[0.010, 0.031]	<.001
Days Between Intervention Launch and Participant Opt-in	-0.001	[-0.002, -0.001]	<.001	-0.012	[-0.015, -0.010]	<.001	-0.001	[-0.001, -0.000]	.001	-0.000	[-0.000, 0.000]	.806
Observations		29810			29810			29810			29810	
Adjusted R^2		0.009			0.006			0.003			0.001	

Note. This table reports the results of four ordinary least squares (OLS) regression models predicting self-reported adherence to public health guidelines for each of the four behavioral questions collected during the intervention among Study 2 participants. Comparisons between the *Q&A* and *direct statement* conditions are derived from Wald tests, as preregistered. In each case, as preregistered, missing values are replaced with the "worst" possible response (i.e., maximal non-compliance). The first regression model predicts whether a given Study 2 participant reported always wearing a mask when leaving home (Model 1, preregistered). The second regression model predicts how many days in the past week a given Study 2 participant reported leaving home for non-essential reasons, reverse-scored such that "0" represents leaving home for non-essential reasons every day and "7" represents never leaving home for non-essential reasons (Model 2, preregistered). The third regression model predicts whether a given Study 2 participant reported washing their hands thoroughly the last time they returned from an outing (Model 3, preregistered). The fourth regression model predicts whether a given Study 2 participant reported touching, hugging, or shaking hands with anyone outside their household, reverse-scored such that "1" represents not touching anyone outside their household, and "0" represents touching someone outside their household (Model 4, preregistered). All models show the main effect of assignment to the Q&A condition. The models also include a continuous control for the day the participant was recruited and an indicator for assignment to the *no-treatment control* condition (not shown here; see Supplement Table 12 for results comparing the *no-treatment control* condition to the *information provision* condition).

Table 5. Regression-Estimated Effects of Q&A Treatment Relative to Direct Statements on Composite Information Seeking and Information Sharing in Study 2.

	Model 1 *Outcome: Information Seeking Composite*			**Model 2** *Outcome: Topic-Specific Information Seeking Composite (Exploratory)*			**Model 3** *Outcome: General Information Seeking Composite (Exploratory)*			**Model 4** *Outcome: Number of phone numbers shared*		
	b	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**
Q&A Condition	0.010	[-0.012, 0.031]	.378	0.032	[0.018, 0.045]	<.001	-0.022	[-0.034, -0.011]	<.001	-0.001	[-0.006, 0.004]	.677
Days Between Intervention Launch and Participant Opt-in	0.000	[-0.000, 0.001]	.522	0.001	[0.000, 0.001]	<.001	-0.001	[-0.001, -0.000]	<.001	-0.000	[-0.000, -0.000]	.046
Observations		29810			29810			29810			29810	
Adjusted R^2		0.007			0.005			0.008			0.000	

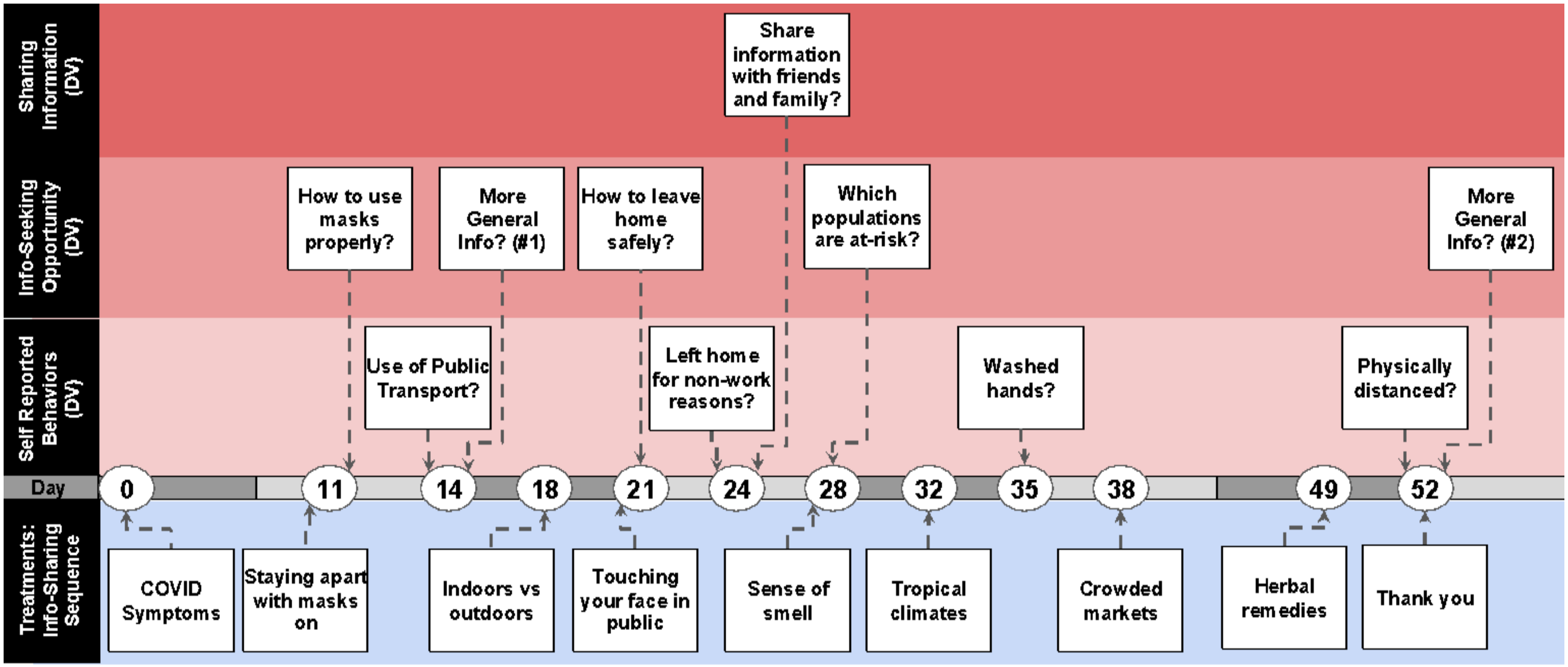
Note. This table reports the results of four ordinary least squares (OLS) regression models predicting information seeking and information sharing among Study 2 participants. Comparisons between the *Q&A* and *direct statement* conditions are derived from Wald tests, as preregistered. The first regression model predicts a given Study 2 participant's information seeking composite score, generated by summing the number of times (out of 5) the participant responded affirmatively to an information seeking opportunity (Model 1, preregistered). The second regression model predicts a given Study 2 participant's topic-specific information seeking composite score, generated by summing the number of times (out of 3) the participant responded affirmatively to an offer of more information about a specific topic related to COVID-19 (Model 2, exploratory). The third regression model predicts a given Study 2 participant's general information seeking composite score, generated by summing the number of times (out of 2) the participant responded affirmatively to an opportunity to access resources with general information about COVID-19 (Model 3, exploratory). The fourth regression model predicts the number of phone numbers of people they would like to keep informed about COVID-19 a given participant in Study 2 shared (Model 4, preregistered). All models show the main effect of assignment to the Q&A condition. The models also include a continuous control for the day the participant was recruited and an indicator for assignment to the *no-treatment control* condition (not shown here; see Supplement Table 16 for results comparing the *no-treatment control* condition to the *information provision* condition).

Table 6. Regression-Estimated Effects of Q&A Ad on Unique Link Clicks in Study 3.

	Model 1 *Outcome: Number of unique clicks on the link to the CDC website included in the ad on a given day* Time period considered: Full experiment			**Model 2** *Outcome: Number of unique clicks on the link to the CDC website included in the ad on a given day* Time period considered: June 18 – July 5		
	b	**95% CI**	**p**	**b**	**95% CI**	**p**
Q&A Ad	0.087	[0.058, 0.116]	<.001	0.107	[0.079, 0.136]	<.001
Dollars Spent on Ad Placement (Log)	3.691	[2.911, 4.472]	<.001	4.924	[-3.598, 13.45]	0.045
Fixed Effects for Day		Yes			Yes	
Observations		52			36	
Adjusted R^2		0.999			0.9448	

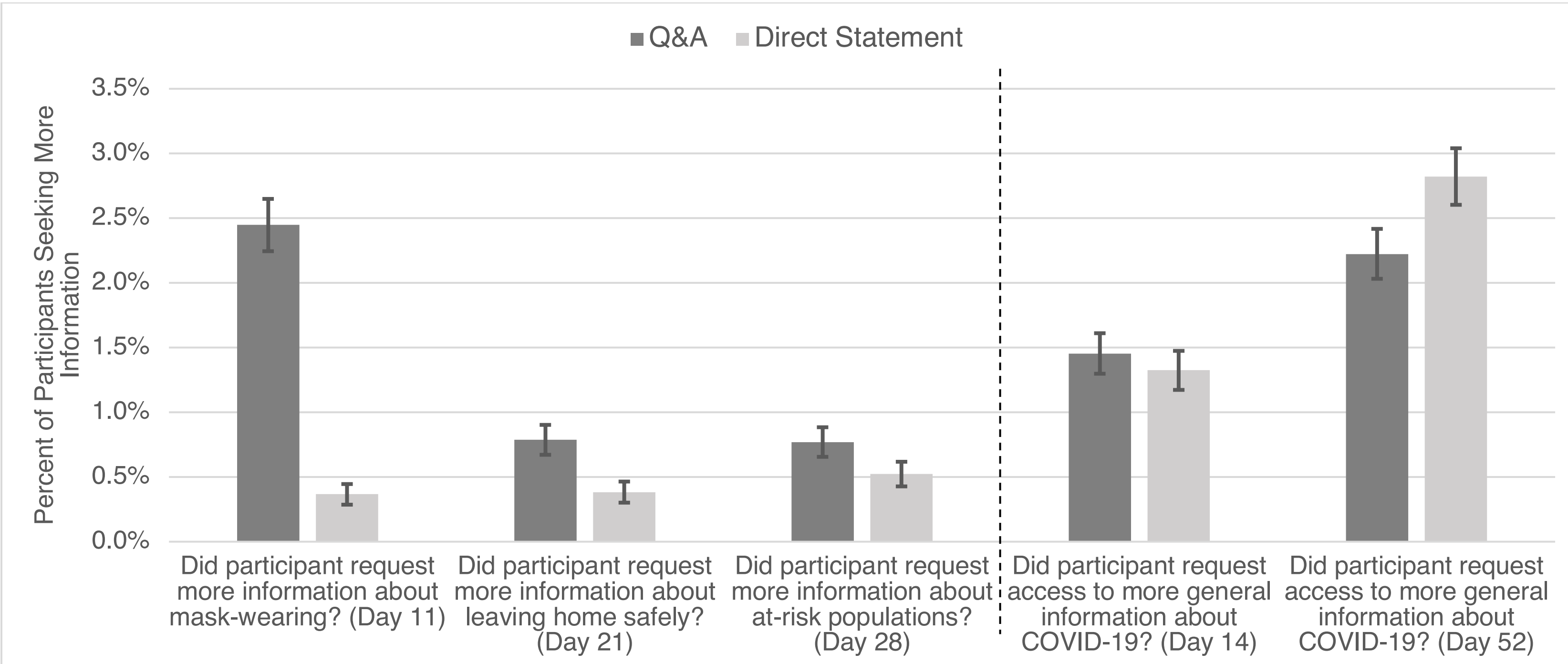
Note. This table reports the results of two log-linear regression models with robust standard errors predicting the effect of the *Q&A* ad relative to the *direct statement* ad on daily unique link clicks. For both models, the dependent variable is the number of unique clicks on the link to the CDC vaccine information page generated by a given ad on a given day. Predictor variables include an indicator for the *Q&A* ad, the (logged) dollars spent bidding on a given ad on a given day, and fixed effects for the day. Model 1 shows the results of this analysis for the full dataset, while Model 2 shows the results of this analysis after the burn-in period only.

Figure 1. Timeline of text messages sent to Study 1 participants in Ghana.



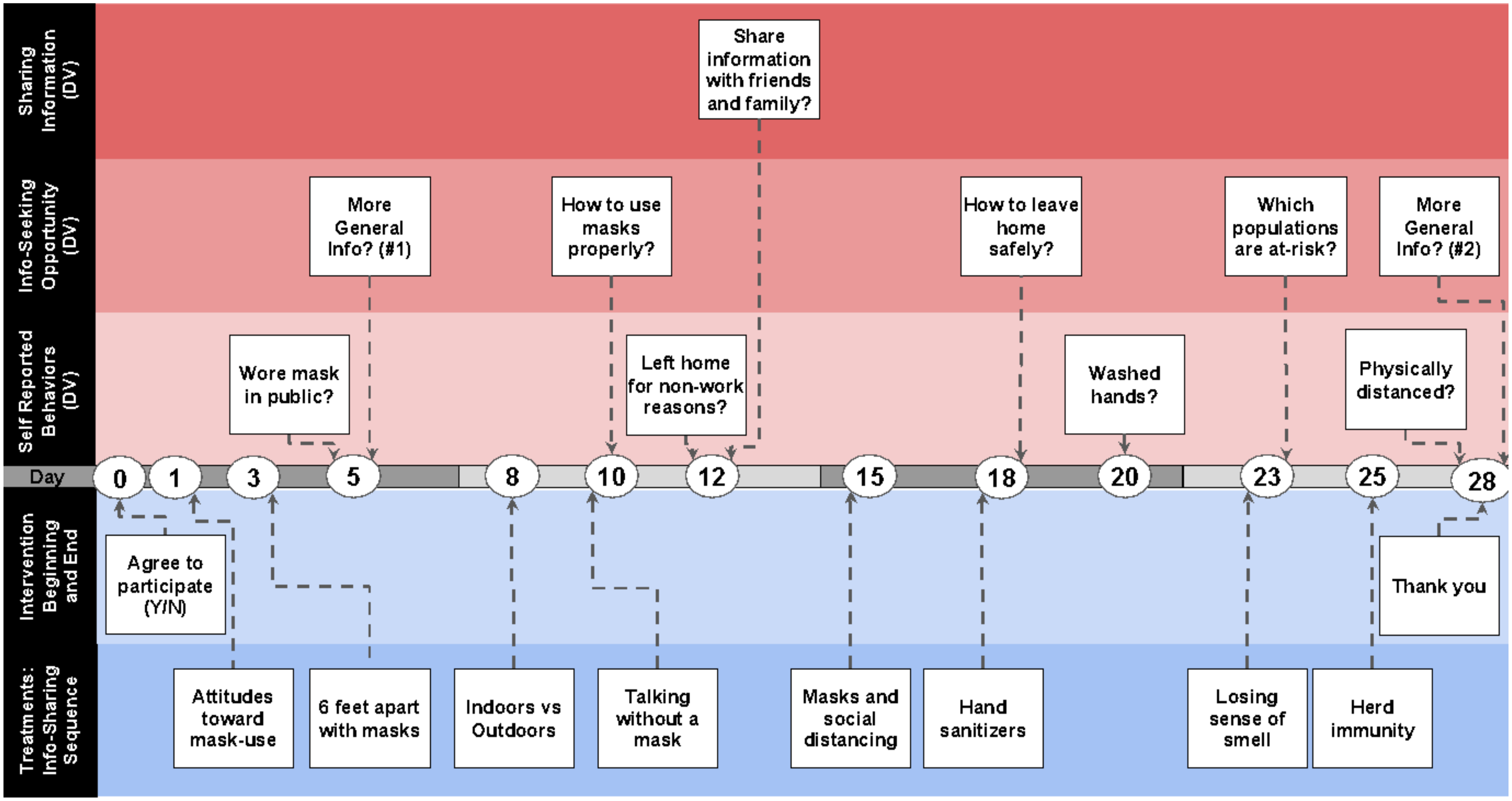
Note. This figure depicts the timing of all text messages sent during the intervention period in Ghana. All participants were recruited between May 22, 2020 and December 1, 2020 via IVR (see Section 2.1) and those who opted in began receiving messages on December 10, 2020 (Day 0). Different types of message sequences are denoted in different colors: the information sharing messages, which vary across treatments, are denoted in light blue, while the dependent variables we collected are in shades of red. The final "Thank you" message is an end-of-study message that is identical across conditions. The arrows denote ordering of message sequences; for example, on day 28, the information-sharing sequence about sense of smell preceded the information-seeking opportunity about at-risk populations. Although we did collect self-reported behaviors, response rates were too low for us to make meaningful inferences from this data (see Footnote 5).

Figure 2. Information Seeking Across Conditions in Study 1



Note. This figure depicts Ghanaian citizens' interest in seeking more information about COVID-19 in Study 1. The dark grey bars represent the proportion of participants who requested more information in the *Q&A* condition while the light grey bars represent the proportion of participants who requested more information in the *direct statement* condition. The three sets of bars on the left of the dotted line depict participants' interest in learning more specific information about masks, leaving home safely, and at-risk populations while the two sets of bars on the right of the dotted line represent participants' interest in learning more general information about COVID-19. Standard error bars are depicted around each proportion.

Figure 3. Timeline of text messages sent to Study 2 participants in Michigan.



Note. This figure depicts the timing of all text messages sent during the intervention period in Michigan. All participants were recruited between November 12, 2020 and December 30, 2020 via text message (see Section 3.1) and those who opted in began receiving messages at 9am the next day (Day 1). As a result, the day of week and day of month on which participants started the intervention depended on the date on which they opted in. Different types of message sequences are denoted in different colors. The information sharing message sequences, which vary across treatments and are not received by participants in the *no-treatment control* condition, are denoted in darker blue, while the dependent variables we collected are in shades of red. The arrows denote ordering of message sequences; for example, on day 18, the information-sharing sequence about hand sanitizers preceded the information-seeking opportunity about how to leave home safely.

Figure 4. Example information-sharing message sequence sent in the Q&A condition (left panel) and direct statement condition (right panel) in Study 2.

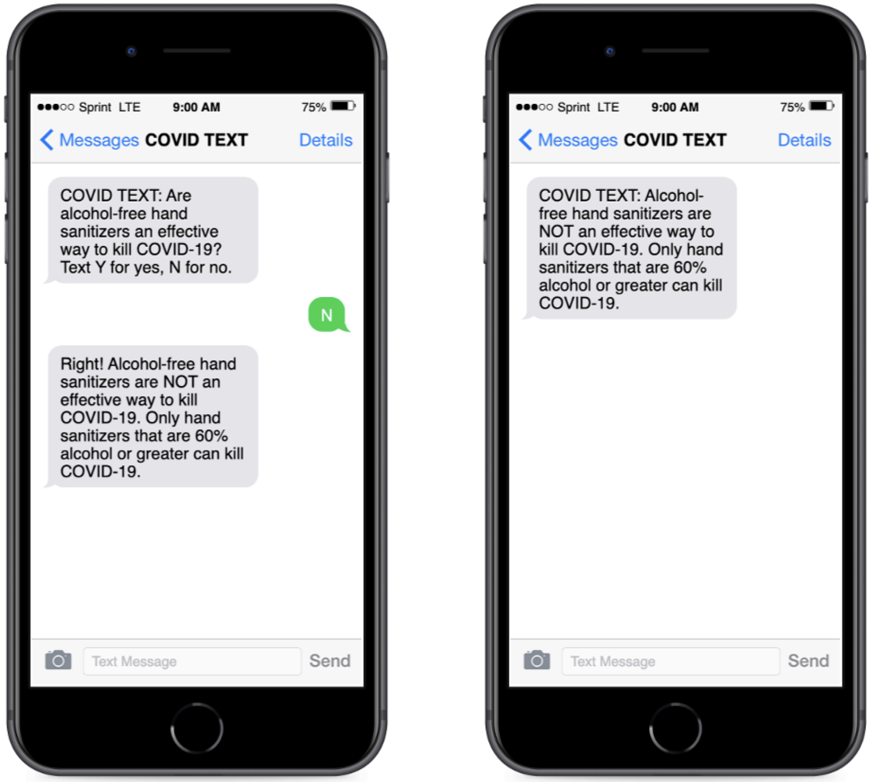
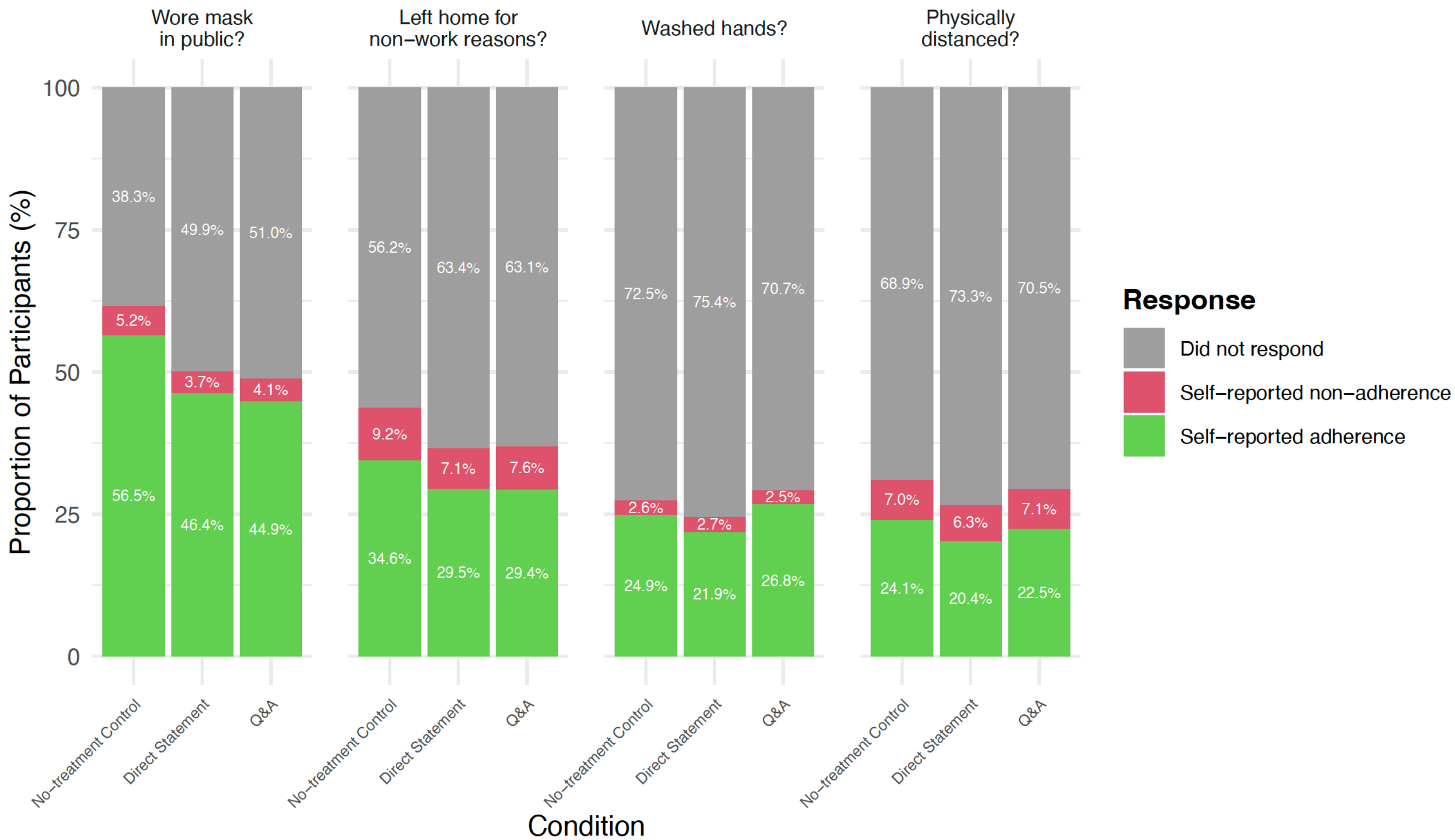
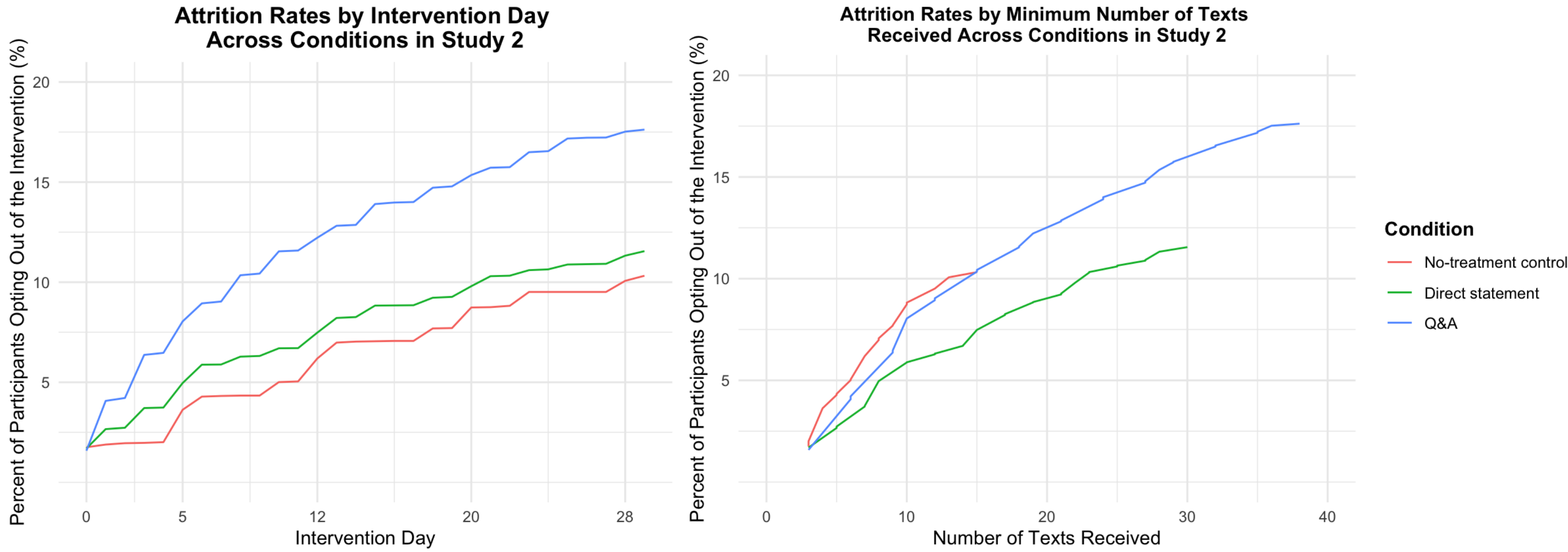


Figure 5. Self-Reported Compliance, Non-Compliance, and Non-Response Rates Across Conditions in Response to Behavioral Measures in Study 2



Note. This figure depicts Michigan residents' responses to self-reported behavioral measures in Study 2. Participants were asked about whether they wore a mask the last time they left home, whether they had left home for unnecessary social outings, whether they had washed their hands the last time they returned home from an outing, and whether they had touched or hugged anyone outside their household during the past week. Each bar represents the responses to one of these four questions (see top labels) across each of the three experimental conditions (see bottom labels). Each bar adds up to 100% and depicts the proportion of participants in the experimental condition who reported they complied with public health recommendations (in green), the proportion of participants in the experimental condition who reported they did not comply with public health recommendations (in red), and the proportion of participants in the experimental condition who did not respond (in grey).

Figure 6. Attrition Rates Across Conditions Over Time (Panel 1) and By Minimum Number of Texts Received (Panel 2) in Study 2



Note. This figure depicts Michigan the rate at which Michigan residents opted out of the text-messaging intervention in Study 2. Across both panels, red lines represent the proportion of participants opting out of the intervention in the *no-treatment control* condition while green lines and blue lines represent the *direct statement* and *Q&A* conditions, respectively. Panel 1 on the left depicts the cumulative proportion of participants who opted out on each of the 29 days of the intervention. Panel 2 on the right depicts the cumulative proportion of participants who had opted out based on the minimum number of texts they had received by that point in the intervention.

Figure 7. Stills of the dynamic Facebook ads the Q&A condition (left panel) and direct statement condition (right panel) in Study 3.

Note. The figure depicts stills of the short videos used in the Facebook ads in Study 3 in the *Q&A* condition (on the left) and *direct statement* condition (on the right). The videos were three-second GIFs in which the text in the ad dynamically popped up. We used videos rather than photographs because they tend to be more eye-catching and are recommended for social media ads. Both ads included a link with the text "click here to learn more about vaccines", which directed users to the CDC's "Key Things to Know About COVID-19 Vaccines" page.