Improving compliance with COVID-19 guidance: a workplace field experiment

Arroyos-Calvera, Danae; Drouvelis, Michalis; Lohse, Johannes; Mcdonald, Rebecca

DOI:
10.1017/bpp.2023.31

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Arroyos-Calvera, D, Drouvelis, M, Lohse, J & Mcdonald, R 2023, 'Improving compliance with COVID-19 guidance: a workplace field experiment', *Behavioural Public Policy*, pp. 1-23. https://doi.org/10.1017/bpp.2023.31

Behavioural Public Policy (2023), 1–23
doi:10.1017/bpp.2023.31

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Improving compliance with COVID-19 guidance: a workplace field experiment

Danae Arroyos-Calvera[1], Michalis Drouvelis[1,2], Johannes Lohse[1] and Rebecca McDonald[1]

[1]Department of Economics, University House, University of Birmingham, Birmingham, UK and [2]CESifo, Munich, Germany
Corresponding author: Michalis Drouvelis, Email: m.drouvelis@bham.ac.uk

(Received 1 December 2022; revised 29 June 2023; accepted 1 September 2023)

Abstract
Compliance with hygiene and other safety measures in the workplace was an important component of society's strategy for reducing infections at the onset of the COVID-19 pandemic, in particular before vaccinations were widely available. We report the results of a field trial of well-established behavioural interventions (social norms, pledging and messenger effects) we implemented to improve compliance with such measures in an occupational setting. We use daily reports of own and other's behaviour to assess the effects of these interventions and supplement these subjective (self-reported) measures with objective data on hand sanitiser usage. The behavioural interventions tested have statistically significant but quantitatively moderate effects on subjective compliance measures and minimal effects on hand sanitiser usage. All effects of our interventions are short-term in nature and dissipate shortly after implementation. Our findings thus provide at most weak support for the notion that typical behavioural interventions can help support compliance with infection prevention measures in the workplace.

JEL Codes: C39; D91

Keywords: social norms; pledge; field experiment; COVID-19; pandemic; workplace safety

Introduction

Prior to the wide availability of vaccines against COVID-19, the reduction of physical contact ("social distancing") was one of the key measures that public health experts recommended when trying to control the spread of the virus within the population. Following this and other recommendations often required large-scale behaviour change in private and professional settings (Van Bavel *et al.*, 2020). In light of the emergence of new variants and reduced vaccine efficacy against these variants (in particular in terms of transmission (see e.g., Andrews *et al.*, 2021), employers may continue using some behavioural interventions to reduce the risk of transmission in occupational settings.

Despite their importance, complying with COVID-19 guidelines in a workplace environment may entail significant private costs, both in terms of productivity and

employee well-being. To avoid these costs, employees may not fully comply with COVID-19 safety measures at work, particularly when they perceive their own risk of contracting the disease to be low or believe they would be unlikely to suffer severe symptoms. If people believe their private compliance costs outweigh their private benefits from avoiding the illness, this could lead to compliance levels below the social optimum (Bos *et al.*, 2020). Moreover, individual compliance may be reduced by fatigue, habits, misconceptions or inattention. Strict rules could force high compliance, but in practice, they may be difficult to enforce, given imperfect monitoring. In the absence of other enforcement, behavioural interventions may help maintain high compliance with social distancing rules in the workplace and beyond (Van Bavel *et al.*, 2020).

Experimental economists have devised a variety of potential interventions that could increase compliance in settings that require voluntary cooperation. We consider three of them. The first is to influence the perception of social norms surrounding compliance (e.g., Bicchieri, 2005; Croson and Shang, 2008; Croson *et al.*, 2009; Gaechter *et al.*, 2013; Goeschl *et al.*, 2018; Bilancini *et al.*, 2020). We implemented this intervention by including information about an injunctive norm in email communication to employees. In these emails, employees were told that a majority of their colleagues believed that they should follow specific safety procedures. This follows a similar approach to Bilancini *et al.* (2020), who study the effects of norms on reading government COVID19 guidance in a survey experiment. The second is to pledge compliance and thus to make a commitment or promise to comply (e.g., Nagin *et al.*, 2002; Camera *et al.*, 2013; Koessler *et al.*, 2019). Employees were invited via email to do this by signing an online form. On the form, they promised they would comply with all safety guidelines. The third is a messenger intervention (Dolan *et al.*, 2012), operationalised through the employee creation of an illustrated COVID-19 safety message. After the safety message was created, it was publicly displayed at the workplace to remind employees of the importance of different safety procedures. Further details about our interventions are provided in section 2.3.

In this paper, we take these three well-established interventions to the field and test their effectiveness in increasing compliance with COVID-19 measures. Our study is set in a workplace that requires employees to move between different offices and machinery on the shop floor. COVID-19 safety measures already in place at the start of the study included maintaining a safe distance from colleagues, abiding by a one-way system and observing good hand hygiene. Over 12 weeks, we implemented the three interventions designed to promote overall compliance with these existing measures.

As the above examples show, these types of interventions have been effective in a variety of settings in the experimental lab and in the field. Various explanations for their effectiveness have been proposed. For instance, individuals may behave in accordance with descriptive or injunctive social norms if they receive dis-utility from contravening established social rules (e.g., Cialdini *et al.*, 1990; Pérez, 2008; Bicchieri and Xiao, 2009). People may change their behaviour in line with their pledge so as to be consistent with themselves (e.g., Bem, 1972; Cialdini, 2009), to avoid dissonance between the promise and their own behaviour (Aronson, 1999) or to avoid guilt (e.g., Kerr and Kaufman-Gilliland, 1994; Charness and

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Dufwenberg, 2006). The messenger intervention, involving illustrated messages created by the participants themselves, should increase compliance because messenger interventions are often more effective when the messenger shares key characteristics with message recipients (Durantini *et al.*, 2006). More generally, the display of message illustrations may increase the salience of existing safety measures, which could address non-compliance resulting from a potential lack of attention or memory (Dolan *et al.*, 2012).

Our contributions are threefold. First, this study belongs to a small set of papers testing behavioural interventions for COVID-19 in a field setting. Notable examples focusing on similar questions include Banerjee *et al.* (2020), who studied the effects of a messenger intervention in a field experiment. The majority of existing studies on COVID-19 are based on framed survey experiments (e.g., Betsch *et al.*, 2020; Bos *et al.*, 2020; Mueller and Rau, 2020), observational studies (e.g., Simonov *et al.*, 2020; Wright *et al.*, 2020) or laboratory experiments (e.g., Branas-Garza *et al.*, 2020; Buso *et al.*, 2020).[1] Second, we collect both self-reported and objective measures of compliance behaviour. Whilst objective compliance data may be considered a gold standard for testing intervention effectiveness, they are typically hard to collect in a non-intrusive way within a field setting. Perceptions of compliance are an important outcome variable in their own right, not only because they allow for conclusions about actual compliance but also because of their association with employees' safety perceptions and hence their well-being at work. Collecting both types of data provides a more comprehensive picture of effectiveness. Third, following the same employees for 12 weeks, we can identify temporal shifts in their perceptions and compliance behaviour.

Implementing our study in a field setting during an ongoing pandemic presented several challenges and limitations that require careful consideration when interpreting our findings. First, the size of our sample was constrained by the relatively small workforce of the organisation we collaborated with. Second, all employees were located at the same on-site location, which prevented us from adopting a between-subjects design due to the high likelihood of treatment spillover. Instead, we rely on a within-subjects design, where interventions were implemented for all employees in-between fallow weeks without interventions. By controlling the timing of the interventions, we can identify their effects compared to the fallow weeks, assuming that their timing did not coincide inadvertently with time-varying factors that could independently influence compliance behaviour, such as local COVID-19 cases or media reporting. In this respect, it is important to note that our study was conducted during the summer of 2020 at a stage of the pandemic where local cases were relatively stable, and the government had lifted the most stringent restrictions on personal movement. In Supplementary Appendix 5, we rule out that the timing of our interventions was correlated with changes in government guidance or spikes in local COVID numbers. We also demonstrate there that our main results are robust to controlling statistically for variables that capture several aspects of the ongoing pandemic dynamics.

Controlling for the temporal dynamics of the COVID pandemic derives its importance from the possibility that economic preferences such as risk and social

[1]Each of these methodologies has its own advantages and disadvantages discussed, for instance, in Harrison and List (2004) or Czibor et al. (2019).

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

preferences are impacted by (local) COVID cases. For instance, Shachat *et al.* (2021) found elevated rates of altruism, cooperation and an increased aversion to risk-taking in the early stages of the COVID-19 pandemic in China. This could be particularly important, since Campos-Mercade *et al.* (2021) show that health behaviours such as those we study here are correlated with pro-sociality. On the contrary, Arroyos-Calvera *et al.* (2023) found that in April 2020, distributional preferences for safety were not significantly different to those elicited more than a year and a half prior to the COVID-19 pandemic outbreak. In this context, it is also essential to highlight that these changes in preferences (when observed) took part in the early stages of the pandemic where uncertainty about its severity was large, while the 10 weeks of our study fell into a period where COVID numbers were more stable and much of the initial uncertainty around COVID had already settled.

Our results suggest that common types of interventions, proven effective in previous investigations, had an overall positive effect on self-reported compliance behaviour. However, even where the effects are statistically significant, effect sizes are moderate. Moreover, the three interventions had no meaningful effect on the objective measure of compliance behaviour (daily use of hand sanitiser). The relatively small effects we observe for self-reported behaviour may reflect low statistical power or a ceiling effect. Most employees already reported that they and their colleagues complied with the guidance most of the time. This accords with recent related findings suggesting that, even prior to the introduction of stringent government regulations, people tended to adapt their behaviour according to their own assessments of the risk they face. That is, voluntary compliance may be higher and more pronounced than individual cost–benefit considerations suggest (Allcott *et al.*, 2020). However, with an infectious disease, there are negative externalities: non-compliance, even by a minority, may significantly harm risk prevention at a group level like a workplace. The impact of research focusing on interventions to change compliance behaviour among this minority of less compliant workers is, therefore, potentially very large.

Design

Setting

The field experiment was conducted at the High Temperature Research Centre (HTRC), a joint endeavour of the University of Birmingham and Rolls-Royce, specialising in production-scale research and experimentation. At the time of the study, amidst the first COVID-19 restrictions, the facility had approximately 120 staff members, with no more than 75 on-site concurrently. Staff duties span across on-site offices and manufacturing/research laboratories, often requiring them to operate in both. Shared facilities include bathrooms, lunchrooms, meeting rooms and corridors, necessitating frequent movement between these areas during a typical workday.

The implementation of the project started in mid-July 2020. At the beginning of that month, English indoor recreational venues such as pubs, restaurants and museums had been allowed to service customers again under heightened hygiene standards, although other businesses such as gyms and beauty salons were still closed to reduce the risk of transmission. In the context of private meetings, people could

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

still only gather indoors if they belonged to the same household, or in groups of up to six people of different households if they were outdoors. Keeping a distance of 1 + m was also mandated in most public and private settings. In workplaces, employers were mandated to provide safety measures.

Prior to the project, the HTRC had already imposed its own comprehensive measures to combat the spread of COVID-19. Measures included 2-m distancing rules aided by floor markings, a one-way system throughout the site, limited meeting room capacities, signage reminding staff about hand hygiene and social distancing, guidance about social distancing during lunch breaks and the provision of hand sanitiser and cleaning wipes. Our research focused on interventions that would increase and maintain compliance with these existing measures, as opposed to the introduction of new ones. One member of our research team visited the HTRC in person in June 2020. On the basis of this visit and in collaboration with the centre's Operational Director, areas of concern, which could plausibly lead to a decline in compliance, were identified.

Procedures

The Operations Director of HTRC introduced the project to employees in a centre-wide "town hall" meeting, being neutral about the specific aims of the project and what would happen. One member of the research team was introduced at an online company meeting. The first activity was an initial questionnaire that elicited baseline estimates of employees' attitudes and behaviour and provided the basis for later interventions (see below), and which elicited demographic information. All communications of the research team with employees were carried out by email, and the questionnaires and relevant emails (including the content of the slide that the Operations Director used to first introduce the project to employees) are reproduced in the Supplementary Appendix.

Randomising the workforce into treatment and control was not feasible, due to the relatively small on-site workforce of around 75 people in the HTRC, and due to the danger of treatment spillover within the single-unit workplace where workers frequently meet and exchange. Therefore, we rolled out our interventions sequentially using a within-subjects design which we describe in more detail below. With the available sample size, we have sufficient power to detect medium-sized effects ($d = 0.5$) at conventional levels of significance and power.[2] For norms-based interventions in the context of COVID-19, a medium-sized effect appears plausible (Goldberg *et al.*, 2020).

With the clear advantages of higher statistical power and practical feasibility of a within-subjects design come several downsides relative to a between-subjects design.[3] Most importantly, identification relies on the assumption that no other factors influence outcome measures in our experiment and particularly that these factors are not correlated with the timing of our interventions. While the timing of each intervention was determined exogenously at the beginning of the experiment by us, it is still

[2]We base this power calculation conservatively on the statistic for a two-sided matched pairs (within-subjects) test with our effective sample size of $N = 57$ (G-power 3.1, $\alpha = 0.05$, $\beta = 0.08$) (Faul *et al.*, 2007). For some of the panel analyses, we perform below power may be larger.

[3]For a general discussion of the relative merits of both approaches, see Charness *et al.* (2012).

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

possible that some observable or unobservable factor coincides with their implementation. In the context of this study, the most important observable factors may be the temporal evolution of COVID-19 cases, public interest in COVID-19 and the stringency of government guidelines. Supplementary Appendix Figure S3 shows that the stringency of regulation and internet search patterns about COVID-19 remained fairly stable during the study period. In terms of COVID-19 case numbers (both locally and in the UK), we see an uptick in the final two weeks of the study. To account for the possibility that this impacts our findings we control for these variables in robustness checks to our main regressions also reported in the Supplementary Appendix.

Each activity (i.e., each questionnaire completed, pledge signed or illustration created) earned points that would be converted to money for charity. This incentive structure is akin to the fairly widespread practice of employees coordinating to fundraise for a charity. It allowed us to incentivise participation in each single element of the study whilst straying away from traditional economic incentives (i.e., linking participation to financial gains) and without being at risk of crowding out individuals' intrinsic motivation (as has been found to happen, for example, with blood donation in Niza *et al.* (2013) and Mellstroem and Johannesson (2008)). The two individuals with the most points at the end of the 12 weeks selected the charities to which the money would be donated. Although participants were not told the exact tariff for each activity, we periodically announced the total earned. In total, £194 was raised for and donated to charity. In addition, to incentivise the elicitation of the social norms and illustration of messages (see below), there were shopping vouchers at stake. Any form of incentivisation (or their absence) may lead to different forms of selection effects. Without incentives, only the most pro-social employees may partake in the study, low monetary incentives may crowd-out intrinsically motivated participants while high monetary incentives may crowd-in the most money-oriented employees. In the end, we decided on a low charitable incentive that may be most attractive to intrinsically motivated employees, also considering the preferences of the partner organisation.

Interventions

We ran three interventions aiming to maintain or increase compliance with the COVID-19 guidelines that were already in place at the partner organisation before the start of the study. We included fallow weeks in which we elicited compliance measures but no interventions took place. These occurred prior to, in between and after the three interventions – i.e., in weeks 1, 4, 7, 11 and 12. These weeks served as a baseline to which intervention weeks were compared. Supplementary Appendix Figure S4 shows the timeline of the study. This design allowed us to observe the short-term effect of interventions in the weeks in which they were introduced, as well as in subsequent weeks when no interventions took place.

Social norms

In weeks 2–3, we implemented the social norms intervention. It closely followed previous work in this literature (Alpizar *et al.*, 2008; Croson and Shang, 2008; Bicchieri and Xiao, 2009) by providing accurate information about an injunctive norm elicited

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

in an initial questionnaire. In contrast to existing studies, we do not compare the effects of providing information on different types of norms (e.g., as in Croson and Shang, 2008). In our context, we expected that there may be larger uncertainty about injunctive than descriptive norms, and therefore we focused our intervention on the former. In the initial questionnaire, we elicited personal and descriptive norms without incentivisation but, for the purpose of our intervention, we elicited injunctive norms in an incentivised fashion using the method introduced in Krupka and Weber (2013).

We elicited these norms surrounding six areas of compliance and identified the two areas where norms differed most between employees ("distancing during lunch-breaks" and "using the one-way system"). While a majority agreed that these rules were important or very important, a significant minority stated that these rules were only moderately important or not important at all. Through communicating the norm perception of the majority we aimed to shift these less compliant norm perceptions and thereby shift behaviour.

We targeted half of the employees in the first intervention week with a daily email message containing the following statements:

> *"The majority of your colleagues believe that you should keep everyone safe by having your lunch at your desk or outdoors." "The majority of your colleagues believe that you should keep everyone safe by sticking to the one-way system."*

In the second intervention week, all employees saw both email statements.

Pledge

In week 5, participants received an invitation to an online form that briefly outlined the most important guidelines. They could sign a pledge to promise to comply with the guidelines at all times. Pledges have been tested previously in the literature and have been shown to have a positive impact on behaviour (e.g., Jacquemet *et al.* 2019; Bazart *et al.* 2022). The pledge was implemented using a Qualtrics form allowing participants to use their cursor to draw a signature. The full pledge text can be found in the Supplementary Appendix. They could complete the pledge once, any time during weeks 5 and 6. Our pledge operated mostly as a voluntary private commitment device, as we did not publicise who had signed it to the workforce.

Messenger

In week 8, participants were invited to submit illustrations on the theme "Protecting Everyone". The theme was based on participants' own reflections of why it is important to follow the guidelines at work, which they were asked to "explain to a colleague" in the initial measurement questionnaire. After analysing their open-text responses to identify themes, other-regarding motives, self-interested motives and protective motives were established, and the "protecting everyone" statement aimed to succinctly reflect all three.

One week later, one illustration was selected by the researchers as the most visually engaging. This picture was prominently displayed at the organisation; the other entries were showcased in a smaller format. The pictures were on display for two

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

weeks (weeks 9–10), serving as reminders of the need to comply with the COVID-19 guidelines. We also included the selected illustration in the daily email communication for two weeks.

To foster participation, the intervention was framed as a contest where the winning entry would receive a £50 shopping voucher, and every entry submitted would accrue points converted into money towards the charity donation.

Outcome Measures

The main outcome measures are drawn from daily "check-in" questionnaires that measured ongoing perceptions of compliance, sent daily across the 12-week experimental period. "Check-in" questionnaires were sent at the same scheduled time at the end of each workday via email. They were held deliberately short to encourage regular participation. For this reason, we asked for compliance with the COVID-19 guidance in general terms instead of more fine-grained questions about specific rules. More precisely, participants reported their own compliance with the COVID-19 guidance and their perceptions of their co-workers' compliance. The other-report is included in case self-reports are biased upwards, either by social desirability bias (Edwards, 1953; Grimm, 2010) or by "illusory superiority" or overconfidence (Svenson, 1981; Van den Steen, 2004). Measures of others' compliance were not intended to be a substitute of own compliance measures, but a complement. Whilst the earlier will be free from the above-mentioned biases, the latter might be vulnerable to a different set of issues (e.g., noise because it is harder to judge others' compliance than our own).

The relevant questions were:

Thinking about work today, please answer these questions:

- I have complied with the COVID-19 guidance at work (7-point scale from "strongly disagree" to "strongly agree")
- My co-workers have complied with the COVID-19 guidance at work (7-point scale from "strongly disagree" to "strongly agree")

In order to mitigate potential demand effects or the desire to appear compliant towards their employer, we made clear to participants that all answers would be anonymous and would only be shared in aggregated form with the operations director or other management. These concerns provide a second reason for eliciting both own compliance and others' compliance. While employees may overstate their own compliance if they (wrongly) assume that their answers may be monitored, there is no strong reason for misrepresenting the compliance of their co-workers.

We also collected an objective measure of compliance: the amount of hand sanitiser used by the workforce per day, by weight. This measurement was conducted by the Operational Director at the end of every workday, across 20 dispensers.

In addition to the daily measures, we captured a variety of relevant attitudes and behaviours at the beginning of the project using a one-off, longer "initial questionnaire" (reported in full in the Supplementary Appendix). It explored the baseline attitudes, understanding and perceived compliance at the HTRC, as well as demographics. Importantly, it provided inputs for the social norms and messenger

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

interventions. Specifically, we elicited the norms that were shown to participants in the social norms treatment, and themes through text responses about the importance of compliance which were summarised in the "Protecting Everyone" statement in the messenger intervention.

Results

Participants

Of the 75 individuals working on-site whose email addresses we could access, 52 completed the initial questionnaire. Table 1 summarises the key demographics and attitudes of our participants. At an average age of 36.8, the workforce at our partner institution is younger than the UK average (41.3) and female employees are in the minority (27%). Twenty-one percent live alone while the remaining 79% live with their partners or families. We use 11 survey items from the scale introduced by Pennycook *et al.* (2020) to assess risk attitudes towards COVID-19. From these items, we construct an index measure between 1 (very low concern) to 7 (very high concern). The majority of employees in our sample are moderately concerned about COVID-19.

We also use the initial questionnaire to elicit information on compliance norms. We elicited views regarding the compliance with six types of behaviours governed by the workplace COVID-19 guidelines (distancing, one-way system, hand hygiene, safe use of bathrooms, safe use of lunch areas, cleaning workspaces). We ask employees how important they personally view compliance with these rules to be (personal norms) using a 1–5 scale where 1 means they consider adhering to the guidelines "not important at all" and 5 means they consider it "extremely important". We also ask respondents how often, on a scale from 0 ("Never") to 100 ("Always"), they believe different groups of people (colleagues, contractors, themselves) follow these guidelines (i.e., a descriptive norm).

We use the methods introduced by Krupka and Weber (2013) to elicit (injunctive) norms about these behaviours in an incentivised way. In particular, we describe settings in which a co-worker breaks a norm around the six main COVID-19 safety guidelines and ask participants how their colleagues would rate such norm breaking on a scale of "very morally inappropriate" (1) to "very morally appropriate" (6). Participants were incentivised for correctly stating the majority view with a chance to win £20 in shopping vouchers. We use these ratings to construct our norms-based intervention. Table 2 shows that all elicited norms paint a similar picture. Participants

Table 1. Key demographics of participants ($n = 52$)

	Mean	SD	Median
Age (years)	36.82	10.53	34.00
Gender (1 = Female)	0.27	0.49	0.00
Single household (1 = Yes)	0.21	0.41	0.00
Corona index (1–7)	4.25	0.75	4.36

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 2. Compliance norm perceptions ($i = 52$)

	1 Distancing	2 One way	3 Hand hygiene	4 Bathroom	5 Lunch breaks	6 Cleaning UP
Personalised norms	4.04	3.60	4.29	4.29	4.25	4.31
Injunctive norms	2.08	2.56	2.06	2.12	2.87	2.13
Descriptive norms (colleague)	73.48	83.69	78.35	88.40	81.92	78.37
Descriptive norms (external)	58.02	73.96	63.21	72.37	67.19	61.60
Descriptive norms (self)	82.52	88.37	82.56	93.87	93.52	87.62

Notes: The personalised norms scale increases in the perceived importance of the rule from 1 (”not important at all”) to 5 (”extremely important”).
The injunctive norm scale increases in moral judgments of breaking a rule from 1 (”very morally inappropriate”) to 6 (”very morally appropriate”).
The descriptive norm scale increases in perceived compliance with the rule, from 0 (”never complies”) to 100 (”always complies”).

believe that a majority abides by all six rules, and own behaviour is deemed more compliant than that of colleagues and external contractors. Most norm violations are considered highly inappropriate and most view abiding by the rules as personally important. Views about the rules around lunch breaks and the one-way system display a slightly higher variance than those related to other behaviours.

Compliance Overview

Figure 1 summarises how self-reported compliance measures change week by week. It averages 1,054 daily self-reports of own and others' compliance behaviour by week. In total, 57 participants did respond to at least one daily questionnaire. However, not all participants completed the daily check-in questionnaire every day. The average participant filled 18.5 questionnaires during the study period, and the maximum number of filled questionnaires by a single participant was 51. The reasons for why participants are not answering the questionnaire on a given day are unknown to us. They may be anything from absence from work, working from home or simply unwillingness to filling in the questionnaire every day. Importantly, participants who did respond less often than the median participant do not report systematically different compliance levels for both outcome measures (Rank-Sum-Test, $p > 0.10$). Participants with below and above median recorded responses also do not differ significantly in any of the demographic attributes shown in Table 1.

Overall, reported compliance is high, averaging at 5.58 (other) and 6.22 (self) on a 7-point scale. Weekly averages reveal three noteworthy patterns. First, employees

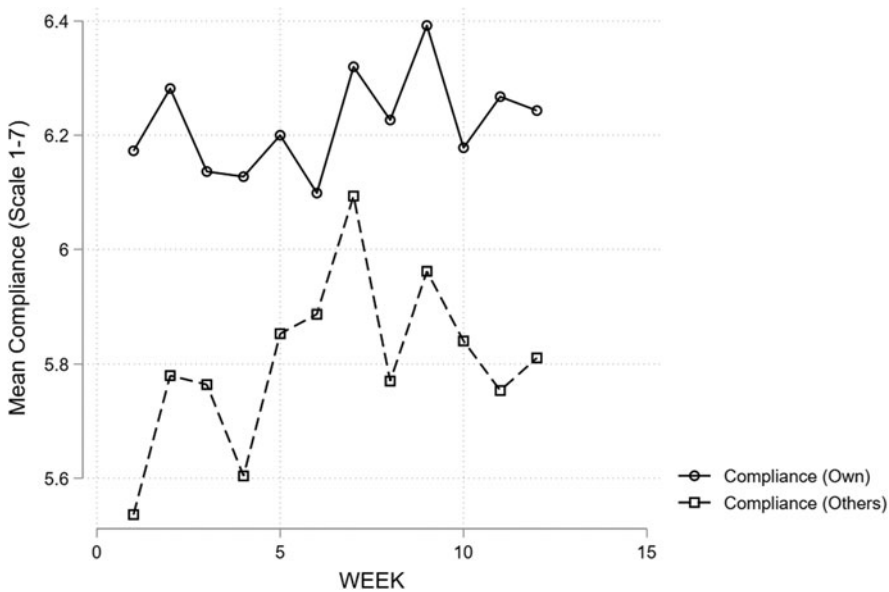

Figure 1. Average compliance rates by week. Reported compliance rates by week. Solid lines show how perceptions of own compliance change over the week. Dashed lines show how perceptions of others' compliance changes over the week.

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

consistently rate others' compliance with COVID-19 guidelines below than their own (Sign-Rank-Test, $p < 0.05$ for all weeks). This could indicate that reports of own behaviour are biased upward, potentially due to social desirability bias (Edwards, 1953; Grimm, 2010) or overconfidence (Svenson, 1981; Van den Steen, 2004). Second, ratings change between intervention weeks (2, 3, 5, 6, 8, 9, 10) and fallow weeks (1, 4, 7, 11, 12). Perceptions of own and others' compliance change synchronously (Spearman's $\rho = 0.61$; $p < 0.001$). Third, there is a weak and positive time trend for others' compliance (Pairwise Correlation $\rho = 0.074$; $p = 0.016$) but not for own compliance (Pairwise Correlation $\rho = 0.038$; $p = 0.21$).

Effect of interventions on self-reported behaviour

Table 3 summarises the effect of interventions on reported compliance by displaying the coefficients of a random effects panel regression. Models (1) and (3) take standardised measures of own compliance as the dependent variable, models (2) and (4) use standardised measures of others' compliance. Taking fallow week 1 as the baseline, we explore the effect of each intervention holding constant the remaining fallow weeks by including additional dummies for each week. Models (3) and (4) account for individual heterogeneity by including fixed effects for each participant. We test whether compliance perceptions across fallow weeks were constant, and this hypothesis cannot be rejected.

The social norms intervention significantly increased own reported compliance by 0.16 standard deviations (SDs). It increased the reported compliance of others by 0.14 SDs, which is weakly significant (Models (1) and (2)). These effects persisted when including individual fixed effects (Models (3) and (4)). The effects of the social norms intervention occurred in the first week of the intervention and weakened in the second. This may suggest that the first intervention derived some of its effectiveness from its novelty rather than from a persistent shift in norms perception. Compliance behaviour in the fallow week directly following the norms intervention (Week 4) did not differ significantly from baseline levels.

The introduction of pledges did not affect employees' own self-reported compliance but did appear to positively influence the compliance behaviour reported for other employees. Reported compliance of co-workers increased by 0.22 SDs without including individual fixed effects and by 0.21 SDs with individual fixed effects. Again, this effect occurred in the first week of the pledging intervention but not in the second. Others' reported compliance behaviour remained elevated during the fallow week directly following the pledge intervention (Week 7), while the level of reported own compliance remained at baseline levels. This may suggest that the pledging intervention had some lasting effect on compliance levels reported for other employees.

The messenger intervention was implemented in two phases. Phase one encompassed the illustration of a message. This phase did not have a statistically detectable effect on own or others' compliance reports. In phase 2, the signs were displayed in the workplace and in email communications. In line with the previous interventions, we found moderately sized positive effects in the first week of the intervention. Displaying the illustrated messages increased reports of own compliance by 0.15

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 3. Effects of interventions

	(1) Own	(2) Others	(3) Own	(4) Others
Norms week 1	0.161** (2.13)	0.138* (1.85)	0.182** (2.31)	0.132* (1.74)
Norms week 2	0.000522 (0.01)	0.185 (1.44)	0.0199 (0.23)	0.188 (1.44)
Pledge Week 1	−0.00134 (−0.01)	0.217** (2.05)	0.00563 (0.05)	0.211* (1.93)
Pledge Week 2	−0.117 (−0.80)	0.177 (1.30)	−0.107 (−0.71)	0.166 (1.18)
Messenger Week 1	−0.122 (−0.86)	0.0648 (0.40)	−0.124 (−0.86)	0.0547 (0.33)
Messenger Week 2	0.151* (1.66)	0.236* (1.89)	0.154 (1.64)	0.221* (1.73)
Messenger Week 3	−0.0171 (−0.15)	0.122 (1.10)	0.00481 (0.04)	0.114 (1.00)
Fallow Week 4	−0.0461 (−0.45)	0.0953 (1.04)	−0.0270 (−0.26)	0.108 (1.11)
Fallow Week 7	0.0152 (0.16)	0.243** (2.32)	0.0199 (0.21)	0.228* (2.10)
Fallow Week 11	−0.0639 (−0.48)	0.0509 (0.30)	−0.0641 (−0.48)	0.0399 (0.23)
Fallow Week 12	−0.0553 (−0.49)	0.0573 (0.35)	−0.0531 (−0.46)	0.0440 (0.26)
Constant	−0.0276 (−0.30)	−0.188 (−1.50)	0.593**** (7.05)	0.612**** (6.07)
Individual fixed effects	No	No	Yes	Yes
Observations	1054	1039	1054	1039
Adjusted R^2	0.003	0.01	0.545	0.538
Prob Chi^2	0.003	0.0759	0.001	0.001

t statistics in parentheses. Random effect regression.
Robust standard errors adjusted for 57 clusters (respondent level).
As we describe in the design section, each intervention was rolled out over multiple weeks. The left-out category is fallow week 1. Norms interventions occurred in weeks 2–3 of the experiment, the Pledge interventions in weeks 5–6 and the Messenger intervention in weeks 8–10.
*p < 0.10, **p < 0.05, ***p < 0.01, ****p < 0.001.

SDs and of others' compliance by 0.24 SDs. These effects weaken when controlling for individual fixed effects, but we retain the significant effect for others' compliance. In the remaining two fallow weeks, behaviour did not differ from the baseline.

Overall, these results suggest that three well-established interventions from experimental economics were partially successful in raising reported compliance behaviour with COVID-19 safety guidelines in a field setting. Discouragingly, their effectiveness is mostly concentrated in the weeks of implementation. When comparing reported

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

compliance levels between fallow week 1 and fallow week 12, the insignificant coefficient in Table 3 indicates that self-reported compliance does not differ between weeks 1 and 12. In other words, the additive effect of all three interventions was not sufficient to create a long-term shift in compliance behaviour, neither for own nor for others' behaviour. Such shifts in behaviour would indicate that our interventions resulted in the formation of habits around COVID-19 avoidance behaviour in the workplace. Habit formation has been observed for other interventions regarding basic hygiene measures like handwashing in rural India (Hussam *et al.*, 2022). Results in Table 3 suggest, however, that the interventions in our study, despite being supported by strong evidence from the economic laboratory, are less well suited to bring about lasting behavioural change in the context of our study. It may well be that this difference is due to different baseline levels of compliance across studies. Reported baseline levels in Hussam *et al.* (2022) are considerably lower than in the context of our study, where self-reported average compliance is already relatively high (5.6 on a 7-point Likert scale) in the week before the first intervention. This generates a possible ceiling effect, as interventions cannot raise compliance further for some workers.

An essential assumption underlying our within-subjects design is that the timing of our interventions is independent of other time-varying factors that could potentially influence compliance behaviour. This assumption ensures that any observed effects can be accurately attributed to the interventions rather than external factors. Although the timing of our interventions was exogenously determined, there remains the possibility of coincidental correlation with other time-varying factors. To address this concern, we conducted additional robustness tests, as shown in Supplementary Appendix Table S6, where we controlled for four factors that could drive reported compliance independently from the interventions we study: National COVID-19 Cases, Local COVID Cases, COVID Rule Stringency and COVID Internet Search Patterns. Each of these variables is a separate indicator of the (perceived) severity of the COVID-19 pandemic and the (government) response to it and hence they may plausibly affect reported compliance in our study.

Our main results are robust to include these variables. Firstly, we found that each of these variables itself had only weak associations with reported compliance behaviour. In this respect, Google Search Trends for the term "COVID19" had the strongest predictive power for compliance behaviour. More importantly, incorporating these additional control variables in our regression analysis did not substantially alter our main treatment effects of the interventions. Especially the effects of the norms interventions are unaltered by including these further control variables. However, it is worth noting that the Pledge intervention ceases to yield a statistically significant effect on either self-reported compliance or compliance of others.

Heterogeneous Treatment Effects

In Table 4, we extend the previous analysis by interacting treatment dummies with two key demographic variables elicited in the initial questionnaire and theoretically relevant to the context of COVID-19 guidance compliance – these are age (relevant as a proxy for vulnerability to the virus) and the degree of COVID-19 concern

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 4. Heterogeneity of own compliance

	Own compliance	(3) Own compliance
Norms Week 1	−0.450 (−1.58)	0.405 (1.38)
Norms Week 2	0.158 (0.61)	0.368 (1.09)
Pledge Week 1	0.0844 (0.35)	0.260 (1.08)
Pledge Week 2	−0.249 (−0.45)	1.272** (2.30)
Messenger Week 1	−0.448 (−1.30)	−0.544 (−0.89)
Messenger Week 2	0.383 (1.41)	0.757* (1.79)
Messenger Week 3	0.242 (0.79)	−0.244 (−0.88)
Fallow Week 4	0.00770 (0.08)	0.00662 (0.07)
Fallow Week 7	0.0573 (0.56)	0.0264 (0.25)
Fallow Week 11	0.00845 (0.06)	0.00832 (0.06)
Fallow Week 12	−0.0715 (−0.59)	−0.0669 (−0.55)
Age (Years)	0.00948 (0.80)	
Age * Norms 1	0.0158* (2.03)	
Age * Norms 2	−0.00235 (−0.37)	
Age * Pledge 1	−0.00366 (−0.52)	
Age * Pledge 2	0.00372 (0.33)	
Age * Messenger 1	0.0111* (1.78)	
Age * Messenger 2	−0.00541 (−0.90)	
Age * Messenger 3	−0.00732 (−0.93)	
Corona Index (1–7)		0.131 (1.24)
Corona * Norms 1		−0.0645 (−0.93)

(*Continued*)

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 4. (*Continued.*)

	Own compliance	(3) Own compliance
Corona * Norms 2		−0.0700 (−0.96)
Corona * Pledge 1		−0.0755 (−1.17)
Corona * Pledge 2		−0.333** (−2.27)
Corona * Messenger 1		0.124 (1.03)
Corona * Messenger 2		−0.133 (−1.41)
Corona * Messenger 3		0.0553 (0.79)
Constant	−0.408 (−0.98)	−0.630 (−1.45)
Observations	905	908

t statistics in parentheses.
*p < 0.10, **p < 0.05, ***p < 0.01, ****p < 0.001.

(relevant as concern is expected to be correlated with adherence to mitigation measures). This allows us to explore the possibility of heterogeneous treatment effects. Such effects would interest those planning to target interventions to specific subgroups. Table 4 contains results for reports about own behaviour. Older participants appear more reactive to the norms and messenger interventions than younger participants. Yet there is little difference in treatment effects between those highly concerned about COVID-19 and those less concerned about it, except that those with higher concern about COVID-19 are less strongly influenced by the pledge intervention. We observe similar qualitative patterns in the reports for others' compliance in Table 5, but the effects are weaker and not statistically significant.

Objective Outcome Measures

Figure 2 displays the temporal evolution of the daily use of hand sanitiser summed over twenty dispensers that were distributed within the workplace. Of course, hand sanitiser usage is not an ideal measure of overall compliance with all safety measures since increased usage may be driven by increases in other risky behaviours such as greater mobility around the site. Moreover, the effectiveness of hand hygiene protocols was increasingly debated in the course of the pandemic. It may thus be the case that some employees did not see a high value in following this workplace guidance. In spite of these considerations, the COVID-19 guidance at the workplace we study clearly suggested observing good hand hygiene. Measuring sanitiser usage hence is an ideal way of capturing compliance with these guidelines. We thus treat the hand sanitiser usage data as a useful complement to the subjective compliance measures that does not suffer from self-reporting. The overall pattern in Figure 2 is of a

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 5. Heterogeneity in others' compliance

	(1)	(4)
	Others' compliance	Others' compliance
Norms Week 1	−0.171 (−0.40)	0.282 (1.00)
Norms Week 2	0.292 (1.18)	0.294 (0.91)
Pledge Week 1	0.401 (1.32)	0.0380 (0.10)
Pledge Week 2	0.0156 (0.06)	0.631** (2.40)
Messenger Week 1	0.154 (0.52)	−0.310 (−0.61)
Messenger Week 2	1.093** (2.19)	0.830* (1.94)
Messenger Week 3	0.318 (1.30)	0.0907 (0.36)
Fallow Week 4	0.169* (1.78)	0.168* (1.80)
Fallow Week 7	0.251** (2.08)	0.252** (2.08)
Fallow Week 11	0.146 (0.84)	0.155 (0.90)
Fallow Week 12	0.142 (0.96)	0.148 (0.99)
Age (Years)	−0.00634 (−0.66)	
Age * Norms 1	0.00845 (0.71)	
Age * Norms 2	−0.000667 (−0.10)	
Age * Pledge 1	−0.00571 (−0.70)	
Age * Pledge 2	0.00555 (0.88)	
Age * Messenger 1	−0.00000809 (−0.00)	
Age * Messenger 2	−0.0223 (−1.58)	
Age * Messenger 3	−0.00483 (−0.76)	
Corona Index (1–7)		−0.0689 (−0.62)

(*Continued*)

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Table 5. (*Continued.*)

	(1) Others' compliance	(4) Others' compliance
Corona * Norms 1		−0.0314 (−0.45)
Corona * Norms 2		−0.00593 (−0.08)
Corona * Pledge 1		0.0345 (0.40)
Corona * Pledge 2		−0.0937 (−1.47)
Corona * Messenger 1		0.112 (1.05)
Corona * Messenger 2		−0.133 (−1.17)
Corona * Messenger 3		0.0132 (0.19)
Constant	−0.00209 (−0.01)	0.0734 (0.17)
Observations	889	892

t statistics in parentheses.
*p < 0.10, **p < 0.05, ***p < 0.01, ****p < 0.001.

positive time trend indicating that hand sanitiser use increased over time: hand sanitiser usage is higher in all weeks (intervention and fallow) after fallow week 0.

The fitted lines highlight differences between intervention and fallow weeks. The fitted relationships are based on a series of regression models (which can be found in Supplementary Appendix Table S7) taking standardised hand sanitiser usage as the dependent variable and investigating whether the difference in usage between intervention and fallow weeks was significant. Overall, sanitiser usage is not significantly different across intervention and fallow weeks.

The time trends depicting usage are significantly flatter in all three intervention weeks, compared to the fallow weeks. This is also demonstrated in the corresponding regression analysis by the interactions between the intervention and the "Day" variable, capturing the time trend. This suggests that the three interventions were to some degree successful in initially shifting hand hygiene upwards but then did not lead to further improvements over time. Once lagged daily sanitiser usage is controlled for, we observe a positive and (weakly) significant effect of the social norms and messenger interventions. We do not observe an effect of the pledge intervention on hand sanitiser use in any model specification.

The overall positive time trend in columns (3) and (4) of Supplementary Appendix Table S7 provides support for the idea that, jointly, our interventions were partially successful in forming and maintaining a habit of hand hygiene in line with the effects reported in Hussam *et al.* (2022).

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

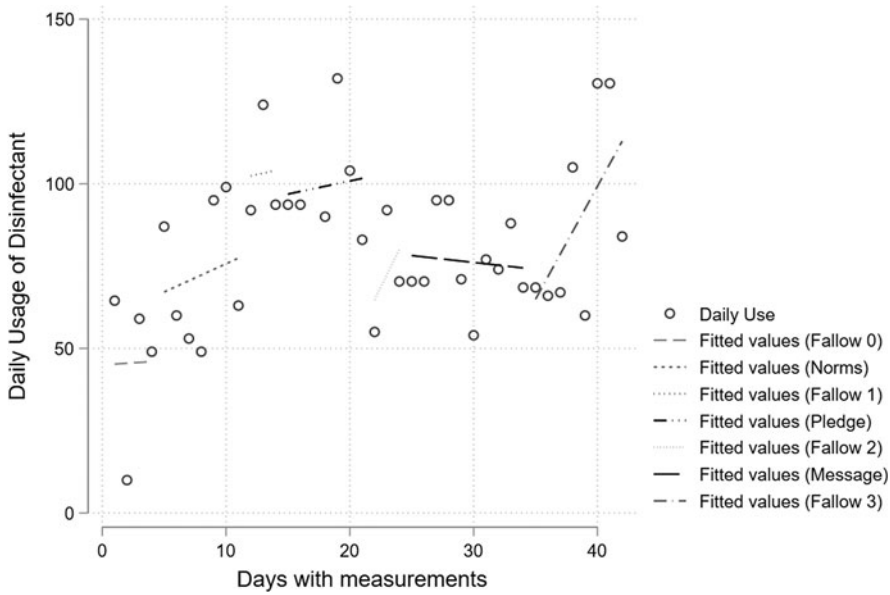

Figure 2. Daily hand sanitiser use. This figure displays the daily use of hand sanitiser over the 41 working days that were covered by our intervention. Fitted lines (linear fit) for each intervention and fallow period.

Discussion

We study the influence of three well-established behavioural interventions on (perceived) compliance with COVID-19 hygiene and other safety measures. We conducted a 12-week intervention study in a workplace setting. Across the 12 weeks, we captured daily measurements of self-reported compliance, as well as an objective measure of hand sanitiser usage.

Our results are mixed. We find some positive effects of all three interventions, with all three associated with statistically significant improvements in perceived compliance in some model specifications. Similarly, we find some evidence of a positive effect of two of the three interventions on hand sanitiser usage, once lags and time trends are controlled for. However, the effects are mostly only significant at the 10% significance level, and in some instances, disappear once individual fixed effects are included in the regression models. Accounting for the fact that we test multiple hypotheses and correcting significance levels accordingly would even suggest that most of our observed effects for the interventions are not statistically significant at corrected levels of significance.

Taken together, this represents a mixed picture of the success of the interventions. Most discouraging is the observation that even when interventions had an effect, this effect was concentrated in the week directly following their implementation. All three interventions together did not shift overall compliance levels at the end of the study period compared to initial compliance levels. This is particularly noteworthy considering that COVID-19 cases were accelerating in the final weeks of the study. To the degree this influences compliance independently from our

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

interventions we would have expected an uptick in compliance in the final two weeks of our study.

The relatively small effects on compliance perceptions may be in part due to a ceiling effect, since most employees reported that they and their co-workers were already mostly adhering to the COVID-19 measures in the workplace, and so there was limited room for improvement on the basis of our interventions. Related findings from the wider literature on COVID-19 support this interpretation, since voluntary compliance with public safety measures has been shown to exceed the levels predicted by cost–benefit analyses on the individual level (Allcott *et al.*, 2020). This finding suggests that successful future compliance nudges may be ones that target those low in compliance, instead of targeting the entire population. Interestingly, we observe a significant gap in reports about others and own compliance. One possible interpretation would be that reports on own compliance are biased upwards because of socially desirable responses, overconfidence or other behavioural biases. To avoid this, focusing on reports about others' compliance may be a preferred option for future studies in similar settings.

Running our interventions at a workplace setting granted us a great opportunity to test three well-established behavioural tools in the field and the potential to achieve a direct positive impact. However, this came at the expense of some of the control that a lab setting offers. Our study was hindered by the available sample size which precluded the randomisation of participants into simultaneous treatment and control groups, instead necessitating a sequential structure with interventions rolled out in turn. A reassuring feature of the results is that compliance reports in the fallow weeks between interventions are indistinguishable, so perceptions of compliance appear to have been reset between interventions.

Of course, whilst this return to baseline compliance in each fallow week is reassuring for our analytical approach, it is less reassuring when considered from the perspective of long-term behaviour change. If these results generalise to other field settings, then it suggests interventions would need to be maintained over time or regularly rotated to maintain the boost to compliance that they deliver when first introduced.

Our field setting offers enhanced external validity compared to survey-based studies or laboratory experiments commonly found in the literature on behavioural interventions and the COVID-19 response. However, it remains uncertain to what extent the specific workplace we examine accurately represents broader workplace environments. The partner organisation, being a relatively small and specialised firm, boasts a highly educated and trained workforce that likely possesses a heightened awareness of COVID-19 risks and a greater responsiveness towards preventive measures. The fact that the behavioural interventions we tested yielded only moderate effects within this population could imply even lower responsiveness to similar interventions in other workplace settings. Conversely, alternative workplace settings may exhibit lower initial compliance levels, thus presenting greater scope for behavioural shifts to materialise.

We believe we are the first to present evidence addressing the effect of behavioural interventions in the workplace applied to COVID-19. From a policy perspective, several insights can be drawn from this study. First, unlike some wider media reporting suggested at the time, individuals in our study setting were highly compliant with and

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

supportive of the COVID-19 measures in place at their workplace and remained so throughout the study period. This aligns well with other literature on this topic which suggests that people were taking voluntary precautions even in the absence of stricter government regulation (Allcott *et al.*, 2020; Hensel *et al.*, 2022). Second, behavioural interventions that have been shown to work well in laboratory conditions or other field contexts may not always be easily transferable to new settings, especially in a dynamically evolving situation such as COVID-19. This is in line with other studies that have found moderate to small effects of common behavioural interventions on compliance with social distancing rules in the context of COVID-19 (Krpan *et al.*, 2021). Third, where we observe an effect of the tested interventions, the effect is rather moderate and does not persist. This is in line with a larger literature that highlights difficulties with creating longer-term behavioural change via short-term interventions (Sunstein, 2017; Nisa *et al.*, 2019). In summary, our results support the view that while behavioural interventions in the workplace setting can support compliance with COVID-19 safety measures, additional interventions may be required to achieve persistently high compliance.

Supplementary Material. To view supplementary material for this article, please visit https://doi.org/10.1017/bpp.2023.31.

Acknowledgements. The authors would like to thank the Editor and two anonymous reviewers for helpful comments and Ganna Pogrebna for setting up the project and for her input to high-level design of the study and text analysis underlying the messenger treatment. Funding from the University of Birmingham is gratefully acknowledged.

References

Allcott, H., L. Boxell, J. C. Conway, B. A. Ferguson, M. Gentzkow and B. Goldman (2020), 'What explains temporal and geographic variation in the early us coronavirus pandemic'? Technical report, National Bureau of Economic Research.

Alpizar, F., F. Carlsson and O. Johansson-Stenman (2008), 'Does context matter more for hypothetical than for actual contributions? Evidence from a natural field experiment', *Experimental Economics*, **11**(3): 299–314.

Andrews, N., J. Stowe, F. Kirsebom, S. Toffa, T. Rickeard, E. Gallagher, C. Gower, M. Kall, N. Groves, A. O'Connell, D. Simons, P. B. Blomquist, A. Zaidi, S. Nash, N. I. B. A. Aziz, S. Thelwall, G. Dabrera, R. Myers, G. Amirthalingam, S. Gharbia, J. C. Barrett, R. Elson, S. N. Ladhani, N. Ferguson, M. Zambon, C. N. J. Campbell, K. Brown, S. Hopkins, M. Chand, M. Ramsay and J. L. Bernal (2021), 'Effectiveness of covid-19 vaccines against the omicron (b. 1.1. 529) variant of concern,' *MedRxiv*.

Aronson, E. (1999), 'The power of self-persuasion', *American Psychologist*, **54**(11): 875.

Arroyos-Calvera, D., J. Covey and R. McDonald (2023), 'Are distributional preferences for safety stable? A longitudinal analysis before and after the covid-19 outbreak', *Social Science & Medicine*, **324**: 115855.

Banerjee, A. V., M. Alsan, E. Breza, A. G. Chandrasekhar, A. Chowdhury, E. Duflo, P. S. Goldsmith-Pinkham and B. Olken (2020), 'Messages on covid-19 prevention in India increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities', NBER working paper (w27496).

Bazart, C., M. Lefebvre and J. Rosaz (2022), 'Promoting socially desirable behaviors through persuasion and commitment: Experimental evidence', *Journal of Behavioral and Experimental Economics*, **101**: 101931.

Bem, D. J. (1972), 'Self-perception theory', *Advances in Experimental Social Psychology*, **6**(1): 1–62.

Betsch, C., L. Korn, P. Sprengholz, L. Felgendreff, S. Eitze, P. Schmid and R. Beohm (2020), 'Social and behavioral consequences of mask policies during the covid-19 pandemic', *Proceedings of the National Academy of Sciences*, **117**(36): 21851–21853.

Bicchieri, C. (2005), *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bicchieri, C. and E. Xiao (2009), 'Do the right thing: but only if others do so', *Journal of Behavioral Decision Making*, **22**(2): 191–208.

Bilancini, E., L. Boncinelli, V. Capraro, T. Celadin and R. Di Paolo (2020), 'The effect of norm-based messages on reading and understanding covid-19 pandemic response governmental rules.' *arXiv preprint arXiv:2005.03998*.

Bos, B., M. A. Drupp, J. Meya and M. Quaas (2020), 'Moral suasion and the private provision of public goods: Evidence from the covid-19 pandemic', *Environmental and Resource Economics*, **76**: 1117–1138.

Branas-Garza, P., D. Jorrat, A. Alfonso, A. M. Espín, T. Garcia Munoz and J. Kovarik (2020), 'Exposure to the COVID-19 pandemic environment and generosity', *Royal Society Open Science*, **9**(1): 210919.

Buso, I. M., S. De Caprariis, D. Di Cagno, L. Ferrari, V. Larocca, F. Marazzi, L. Panaccione and L. Spadoni (2020), 'The effects of covid-19 lockdown on fairness and cooperation: Evidence from a lablike experiment', *Economics Letters*, **196**: 109577.

Camera, G., M. Casari and M. Bigoni (2013), 'Binding promises and cooperation among strangers', *Economics Letters*, **118**(3): 459–461.

Campos-Mercade, P., A. N. Meier, F. H. Schneider and E. Wengström (2021), 'Prosociality predicts health behaviors during the covid-19 pandemic', *Journal of Public Economics*, **195**: 104367.

Charness, G. and M. Dufwenberg (2006), 'Promises and partnership', *Econometrica*, **74**(6): 1579–1601.

Charness, G., U. Gneezy and M. A. Kuhn (2012), 'Experimental methods: Between-subject and within-subject design', *Journal of Economic Behavior & Organization*, **81**(1): 1–8.

Cialdini, R. B. (2009), *Influence: Science and practice*, **Vol. 4**. Boston, MA: Pearson education.

Cialdini, R. B., R. R. Reno and C. A. Kallgren (1990), 'A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places', *Journal of Personality and Social Psychology*, **58**(6): 1015.

Croson, R. and J. Y. Shang (2008), 'The impact of downward social information on contribution decisions', *Experimental Economics*, **11**(3): 221–233.

Croson, R., F. Handy and J. Shang (2009), 'Keeping up with the joneses: The relationship of perceived descriptive social norms, social information, and charitable giving', *Nonprofit Management and Leadership*, **19**(4): 467–489.

Czibor, E., D. Jimenez-Gomez and J. A. List (2019), 'The dozen things experimental economists should do (more of)', *Southern Economic Journal*, **86**(2): 371–432.

Dolan, P., M. Hallsworth, D. Halpern, D. King, R. Metcalfe and I. Vlaev (2012), 'Influencing behaviour: The mindspace way', *Journal of Economic Psychology*, **33**(1): 264–277.

Durantini, M. R., D. Albarracin, A. L. Mitchell, A. N. Earl and J. C. Gillette (2006), 'Conceptualizing the influence of social agents of behavior change: A meta-analysis of the effectiveness of HIV-prevention interventionists for different groups', *Psychological Bulletin*, **132**(2): 212.

Edwards, A. L. (1953), 'The relationship between the judged desirability of a trait and the probability that the trait will be endorsed', *Journal of Applied Psychology*, **37**(2): 90.

Faul, F., E. Erdfelder, A.-G. Lang and A. Buchner (2007), 'G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences', *Behavior Research Methods*, **39**(2): 175–191.

Gaechter, S., D. Nosenzo and M. Sefton (2013), 'Peer effects in prosocial behavior: Social norms or social preferences?', *Journal of the European Economic Association*, **11**(3): 548–573.

Goeschl, T., S. E. Kettner, J. Lohse and C. Schwieren (2018), 'From social information to social norms: Evidence from two experiments on donation behaviour', *Games*, **9**(4): 91.

Goldberg, M. H., A. Gustafson, E. Maibach, S. van der Linden, M. T. Ballew, P. Bergquist, J. Kotcher, J. R. Marlon, S. Rosenthal and A. Leiserowitz (2020), Social norms motivate covid-19 preventive behaviors. Mimeo.

Grimm, P. (2010), *Social desirability bias*. Chichester, UK: Wiley international encyclopedia of marketing.

Harrison, G. W. and J. A. List (2004), 'Field experiments', *Journal of Economic Literature*, **42**(4): 1009–1055.

Hensel, L., M. Witte, A. S. Caria, T. Fetzer, S. Fiorin, F. Götz, M. Gomez, J. Haushofer, A. Ivchenko, G. Kraft-Todd, E. Reutskaja, C. Roth, E. Yoeli and J. M. Jachimowicz (2022), 'Global behaviors, perceptions, and the emergence of social norms at the onset of the COVID-19 pandemic', *Journal of Economic Behavior and Organization*, **193**: 473–496.

Hussam, R., A. Rabbani, G. Reggiani and N. Rigol (2022), 'Rational habit formation: Experimental evidence from handwashing in India', *American Economic Journal: Applied Economics*, **14**(1): 1–41.

https://doi.org/10.1017/bpp.2023.31 Published online by Cambridge University Press

Jacquemet, N., S. Luchini, J. Rosaz and J. F. Shogren (2019), 'Truth telling under oath', *Management Science*, **65**(1): 426–438.

Kerr, N. L. and C. M. Kaufman-Gilliland (1994), 'Communication, commitment, and cooperation in social dilemma', *Journal of Personality and Social Psychology*, **66**(3): 513.

Koessler, A.-K., B. Torgler, L. P. Feld and B. S. Frey (2019), 'Commitment to pay taxes: Results from field and laboratory experiments', *European Economic Review*, **115**: 78–98.

Krpan, D., F. Makki, N. Saleh, S. I. Brink and H. V. Klauznicer (2021), 'When behavioural science can make a difference in times of covid-19', *Behavioural Public Policy*, **5**(2): 153–179.

Krupka, E. L. and R. A. Weber (2013), 'Identifying social norms using coordination games: Why does dictator game sharing vary?', *Journal of the European Economic Association*, **11**(3): 495–524.

Mellstroem, C. and M. Johannesson (2008), 'Crowding out in blood donation: was Titmuss right?', *Journal of the European Economic Association*, **6**(4): 845–863.

Mueller, S. and H. A. Rau (2020), 'Economic preferences and compliance in the social stress test of the covid-19 crisis', *Journal of Public Economics*, **194**: 104322.

Nagin, D. S., J. B. Rebitzer, S. Sanders and L. J. Taylor (2002), 'Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment', *American Economic Review*, **92**(4): 850–873.

Nisa, C. F., J. J. Bélanger, B. M. Schumpe and D. G. Faller (2019), 'Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change', *Nature Communications*, **10**(1): 4545.

Niza, C., B. Tung and T. M. Marteau (2013), 'Incentivizing blood donation: Systematic review and meta-analysis to test Titmuss' hypotheses', *Health Psychology*, **32**(9): 941.

Pennycook, G., J. McPhetres, B. Bago and D. Rand (2020), Predictors of attitudes and misperceptions about covid-19 in Canada, the UK, and the USA. Mimeo.

Pérez, R. L. (2008), 'Aversion to norm-breaking: A model', *Games and Economic Behavior*, **64**(1): 237–267.

Shachat, J., M. J. Walker and L. Wei (2021), 'How the onset of the covid19 pandemic impacted pro-social behaviour and individual preferences: Experimental evidence from China', *Journal of Economic Behavior & Organization*, **190**: 480–494.

Simonov, A., S. K. Sacher, J.-P. H. Dube and S. Biswas (2020), 'The persuasive effect of fox news: non-compliance with social distancing during the covid-19 pandemic', Technical report, National Bureau of Economic Research.

Sunstein, C. R. (2017), 'Nudges that fail', *Behavioural Public Policy*, **1**(1): 4–25.

Svenson, O. (1981), 'Are we all less risky and more skillful than our fellow drivers?', *Acta Psychologica*, **47**(2): 143–148.

Van Bavel, J. J., K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, J. Drury, O. Dube, N. Ellemers, E. J. Finkel, J. H. Fowler, M. Gelfand, S. Han, S. A. Haslam, J. Jetten, S. Kitayama, D. Mobbs, L. E. Napper, D. J. Packer, G. Pennycook, … R. Willer (2020), 'Using social and behavioural science to support covid-19 pandemic response', *Nature Human Behaviour*, **4**: 460–471.

Van den Steen, E. (2004), 'Rational overoptimism (and other biases)', *American Economic Review*, **94**(4): 1141–1151.

Wright, L., A. Steptoe and D. Fancourt (2020), 'What predicts adherence to covid-19 government guidelines? Longitudinal analyses of 51,000 UK adults', *medRxiv*.

Cite this article: Arroyos-Calvera D, Drouvelis M, Lohse J, McDonald R (2023). Improving compliance with COVID-19 guidance: a workplace field experiment. *Behavioural Public Policy* 1–23. https://doi.org/10.1017/bpp.2023.31