Behavioural Public Policy (2026), **10**, 292–316
doi:10.1017/bpp.2023.33

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The effect of timers and precommitments on handwashing: a randomised controlled trial in a kitchen laboratory

James Thom[1], Sarah Bowen[1], Yuchen Yang[1], Sanjeev Devarajan[1], Helen Doran[1], Marios Zampetis[1], Trisevgeni Papakonstantinou[2], Maria McDonagh[3], Robert McPhedran[1], Ben Toombs[1], Ayla Ibrahimi Jarchlo[4], Alice Rayner[4], Philip Jones[4] and Natalie Gold[1,5]

[1]Behavioural Practice, Verian, 4 Millbank, London SW1P 3JA, UK, [2]Department of Experimental Psychology, University College London, London WC1E 6BT, UK, [3]Department of Economics, University of Birmingham, Birmingham B15 2TT, UK, [4]Social Science Team, Food Standards Agency, Clive House, 70 Petty France, London SW1H 9EX, UK and [5]Centre for Philosophy of Natural and Social Science (CPNSS), London School of Economics London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK
Corresponding author: Natalie Gold, email: natalie.gold@veriangroup.com

(Received 28 November 2022; revised 17 October 2023; accepted 19 October 2023)

Abstract

Many foodborne illness outbreaks originate in food service establishments. We tested two behavioural interventions designed to improve the duration and quality of handwashing. We ran a three-armed parallel trial in a laboratory kitchen, from 7 March to 27 May 2022. Participants were $n = 195$ workers who handle food. We randomly allocated participants to three groups: Timer – tap-mounted timer that counted seconds while participants washed their hands; Precommitment – agreed to five statements on good hand hygiene before attending the kitchen; and Control. Participants completed a food preparation task under time pressure. Cameras focused on the sink captured handwashing. Outcome measures were number of times participants washed their hands; number of times they washed their hands using soap; number of times they washed using soap and washed the backs of their hands; and mean duration of handwashing attempts using soap. Participants in Timer washed their hands for 1.9 s longer on average than Control ($\beta = 2.20$, 95% CI = 0.34-4.06, $p = 0.021$). Participants in Precommitment washed their hands for 2.5 s longer on average than Control ($\beta = 2.30$, 95% CI = 0.33-4.27, $p = 0.022$). We found no statistically significant differences on any other outcome measure.

Keywords: Behavioural intervention; food safety; hand hygiene; handwashing; precommitments

Introduction

Background and aims

In 2009 in the UK, there were over 500,000 estimated cases of foodborne disease due to known pathogens (Food Standards Agency, 2014). Foodborne illnesses are caused

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

by eating food contaminated with bacteria, viruses, other parasites, or chemical contaminants like heavy metals. Recent estimates put the number of yearly cases of foodborne illness at 2.4 million in the UK, imposing an estimated total burden of £9 billion (Daniel *et al.*, 2018). Many foodborne illness outbreaks originate in food service establishments, with food workers' poor personal hygiene as a significant contributor to outbreaks (Olsen *et al.*, 2000). For example, eating out accounts for an estimated 37% of all foodborne norovirus cases, and takeaways account for 26% (Food Standards Agency, 2022). Workers' hand hygiene is estimated to account for 89% of variance in outbreaks caused by food contaminated with pathogens in food service establishments (Guzewich and Ross, 1999). A more recent analysis of outbreaks associated with restaurants from 1998 to 2013 found that the most commonly reported contributing factors were those related to food handling and preparation practices in the restaurant (Angelo *et al.*, 2017).

The present study aimed to test the effects of behavioural interventions aimed at improving handwashing behaviour in employees of Food Business Operators (FBOs). In the first stage of the study, we carried out a rapid review of the existing literature on hand hygiene interventions in FBO workers to identify the most promising interventions to test. We then carried out a randomised controlled trial in a kitchen laboratory to provide evidence of their effectiveness.

Review of literature on hand hygiene interventions in food handlers

Poor hand hygiene can be caused by washing for too short a time, not washing frequently enough, or by poor technique – in particular, failing to apply soap to all parts of the hand. Most hand hygiene studies investigate frequency (see, e.g., studies found in (Huis *et al.*, 2012)). However, arguably, frequency of handwashing is insufficient because if hands are not washed for long enough or parts of the hand are missed, then pathogens are still likely to remain (Öncü and Vayısoğlu, 2021). Preliminary research on hand hygiene of UK food handlers, conducted before the onset of the COVID-19 pandemic, identified knowledge gaps around effective handwashing practice in food handlers, including around the length of time required for washing and drying hands (Ipsos Mori, 2017). This pre-pandemic research also found behavioural barriers to handwashing: participants said they were likely to forget to wash when busy or under pressure, that there was little expectation of handwashing from seniors, and necessary materials were often not available.

Post-pandemic, it seems less likely that hand hygiene failures are due to knowledge gaps. One of the UK Government's messages during the COVID-19 pandemic was 'Hands. Face. Space' and well-publicised official guidance was to wash for 20 s (NHS, 2022). In the US, the Centre for Disease Control promoted 'Five Steps to Wash Your Hands the Right Way', one of which was scrubbing hands for 20 s (CDC, 2022). However, it is well-known that knowledge does not on its own lead to action. For instance, in the case of handwashing, a meta-analysis of food safety training in commercial settings found that training improved knowledge in most studies, but it did not lead to any change in handwashing behaviour in five of the six behaviours examined (McFarland *et al.*, 2019). Indeed, reviews of hand hygiene behaviours in various domains have found that the most common aim of

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

interventions is to increase knowledge, with changes to people's motivations and physical environments either underexplored or potentially more effective (Huis *et al.*, 2012; Staniford and Schmidtke, 2020). Behavioural interventions may be able to bridge this 'knowledge-action gap'.

Recent research has shown the potential benefits of interventions on the physical environment to improve duration, with two studies finding that installing timers at sinks were effective at increasing the average duration of handwashing attempts by food handlers. Her *et al.* (2019) installed tap-mounted timers that counted down 30 s when the water was turned on, alongside informational posters about their use and proper handwashing technique. The authors reported a statistically significant increase in mean handwashing duration from a pre-treatment baseline (11.6–15.7 s). Similarly, Yu *et al.* (2018) found statistically significant increases in the frequency (71.2–102.7 times/day), mean duration (12.8–19.2 s) and the quality of handwashing attempts – relative to a pre-treatment baseline – when they introduced a 'motivational' soap dispenser that played 18 s of music whenever food handlers applied soap.

In terms of intervening on motivations, few studies have attempted to improve hand hygiene by intervening on intentions (Huis *et al.*, 2012); however, some researchers have recommended that people use 'implementation intentions' to increase handwashing (Michie *et al.*, 2020). In other domains, the act of making a precommitment or forming a plan to translate intention into action has been shown to support the regulation of existing behaviour and the uptake of a new behaviour. Forming specific implementation intentions of the form 'Whenever situation *x* arises, I will initiate the goal-directed response *y*!' (Gollwitzer, 1999; Brandstätter *et al.*, 2001) has been shown to be an effective strategy when tackling self-regulatory problems (e.g., remembering to act, seizing opportunities and overcoming initial reluctance) (Gollwitzer and Sheeran, 2006). Additionally, previous research has demonstrated the effectiveness of committing to action in advance on the uptake of a wide range of behaviours, including weight loss (Coupe *et al.*, 2019), smoking cessation (Black *et al.*, 2020) and safe water consumption (Inauen and Mosler, 2016). In principle, if someone has the goal of washing their hands appropriately, then making precommitments or forming implementation intentions could help them achieve their goal. However, we are not aware of any studies exploring the effects of these strategies on handwashing.

Two issues that have beset previous research in this area are experimental design and outcome measures, with most studies using observational pre–post designs or self-reported measures (McFarland *et al.*, 2019; Insfran-Rivarola *et al.*, 2020). It is difficult to run a gold-standard randomised controlled trial in a field setting because there are usually only a small number of kitchens (or bathrooms or hospital wards) available to study, so there are not enough units for a well-powered trial. This is compounded by problems of outcome measurement, with the most reliable outcome measures also being the most labour-intensive (Haas and Larson, 2007).

Outcome measures pose particularly intractable problems, requiring trade-offs to be made. Self-reported measures may be a cheap way of gathering a lot of data (Ainsworth *et al.*, 2021), but they tend to be limited to measures of handwashing frequency. In addition, there is a general problem of using self-reports to measure a socially desirable behaviour and some more specific problems around people's answers being sensitive to the way that the survey question is written (Hansen

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

et al., 2021). One way that researchers have tried to avoid self-reports is by measuring soap use as a proxy for handwashing (e.g. (Judah *et al.*, 2009; Porzig-Drummond *et al.*, 2009)). This tends to be done in studies in bathrooms; researchers measure soap use over a period of time whilst also counting the number of people entering into the bathroom, in order to get an average measure of soap use per head. Of course, this method is unable to distinguish a few people using a lot of extra soap from an effect that is more evenly spread across the population. It also does not provide information about duration or quality of handwashing (Yu *et al.*, 2018).

Another way to avoid self-reports is to use direct observation. This allows for the most accurate determination of duration and quality. However, as well as being time- and resource-intensive, there is the possibility that being aware of being observed can itself influence behaviour: the 'Hawthorne Effect' (Haas and Larson, 2007; Gould *et al.*, 2017a, 2017b). Therefore, direct observation seems to be rather rare in the literature, but even when there is direct observation, it may be operationalised as a frequency checklist (e.g. (Husain *et al.*, 2016; Štefančič and Jevšnik, 2020)). Adjustments that may decrease the likelihood of the Hawthorne Effect include the use of confederates, who pretend not to be watching (Hansen *et al.*, 2021). However, this raises ethical issues around deception and may result in the confederate not being close enough to make detailed observations and therefore only providing a measure of handwashing frequency. Cameras provide a less labour-intensive and slightly more indirect measure of observation, which balance the ability to get accurate measures with a slightly lower feeling of being watched (Chapman *et al.*, 2010; Birnbach *et al.*, 2013; do Prado *et al.*, 2015; Yu *et al.*, 2018; Her *et al.*, 2019; Barrett and Feng, 2020).

These observations are congruent with the outcome of a recent Cochrane review of interventions to increase hand hygiene compliance in hospitals, which found that it is difficult to draw many conclusions from the literature because of twin problems of a lack of methodological rigour and lack of concealment of the intervention (Gould *et al.*, 2017a, 2017b; Gould *et al.*, 2018). (Though we also note that hospitals are not cognate settings to food businesses, so apart from having similar methodological problems, that literature is not as pertinent here. We return to this in the discussion.)

Approach and hypotheses

We designed a randomised controlled trial in a kitchen laboratory to test the effect of two behavioural interventions aimed at improving handwashing behaviour in employees of FBOs: timers and precommitments. The use of a laboratory setting allowed for individual randomisation, with the outcome measures of frequency, duration and technique being recorded on video cameras. Participants were given a burger-making task under time pressure, to mimic the experience of being busy or feeling under pressure, which are cited as a barrier to handwashing in restaurant kitchens. There were decoy cameras on the work surfaces, as well as ones taking recordings by the sink, to prevent participants guessing what the purpose of the laboratory trial was and to minimise the experimenter demand effect.

Hypothesis 1: Installing a timer at the sink and a poster prompting proper handwashing technique will increase the duration and quality of handwashing attempts

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

relative to control. Both the timer and poster are concerned with how food handlers wash their hands, not how often they do so, so we hypothesised no effect on handwashing frequency. Frequency was nonetheless of interest, as a reduction in handwashing frequency was a plausible 'backfire' effect if food handlers find handwashing for the full 20 s unduly burdensome.

Hypothesis 2: Asking participants to make a series of pre-trial commitments to comply with specific elements of the FSA's guidance on duration, frequency and quality of handwashing will cause an increase in measures of all three, relative to control.

Methods

Trial design

The trial used a three-arm between-subjects design. We carried out testing in a laboratory kitchen: a cookery school with three separate workstations, screened off from each other. Up to three participants were tested at a time, and each was assigned their own workstation and sink for use (see Figure 1). (The cookery school was in Wandsworth, London, UK. https://theavenuecookeryschool.com/contact/)

We asked all participants to complete a food preparation task following a set recipe. The recipe required that participants handle raw meat and included multiple steps at which handwashing would be expected (eight handwashing occasions expected in total, if the task was completed). The task instructions emphasised the need to work as quickly as possible because food handlers say they are more likely to forget to wash their hands during busy periods (Ipsos Mori, 2017). Before attending the laboratory session, all participants were required to complete a short online training session and post-training quiz, in their own time.

At recruitment, we randomly assigned participants with equal probability to one of the three treatment groups:

In the Timer group, two interventions were placed at the sink: (i) a tap-mounted timer that counts seconds while participants are washing their hands and (ii) a poster encouraging participants to use it to time 20 s rubbing their hands with soap, and showing the six steps of effective handwashing technique from the FSA's Safer food, better business guidance (Food Standards Agency, 2019). (Figure. 2 shows the poster, which includes an image of the timer device.)

In the Precommitment group, the pre-session training included a prompt to commit to a series of statements about handwashing – adapted from FSA guidelines – by electronically signing their name (typing it into a textbox). The statements were:

- *If I touch raw meat, then I will wash my hands afterwards*
- *If I am going to touch or handle any food, then I will wash my hands first*
- *When I wash my hands, I will wet my hands under warm running water and use soap*
- *When I wash my hands, I will wash the back of each hand by rubbing against my palms*
- *When I wash my hands, I will wash my fingertips by rubbing against my other hand's palm*

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press



Figure 1. Image of workstation in the experimental laboratory kitchen.

Posters reminding *Precommitment* group participants of these handwash statements were displayed to participants in view of the sink and the workstation during the food preparation task (see Figure. 3).

Participants in the *Control* group received no handwashing intervention other than the baseline training completed by all participants. Otherwise, all environmental influences were held constant.

Each laboratory session consisted of 3 slots and each slot was quasi-randomly assigned to a different treatment group (for each time and date, the three slots were sorted into a random order using Python, with the function random.shuffle ()). The recruitment agency then recruited into a session's slots in order. For example, if a session's treatments are shuffled into [2, 3, 1], then the first participant recruited for that session was assigned to Group 2, the second recruited to the control, and the third to treatment 1. To minimise the risk of bias based on time of session, participants attending the same session were assigned to different treatment groups. Likewise, to minimise the risk of differences between the workstations confounding treatment effects, the treatment groups assigned to each workstation were rotated between sessions. Opaque barriers were installed to physically separate participants being tested in the same session and minimise potential spillover effects (i.e., participants washing their hands differently because they see other participants doing so or can see interventions from other treatment groups).

We did not inform participants about the focus of this study, or of the FSA's involvement, until the end of the experiment. When initially consenting to take

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press



Figure 2. Poster placed by the sink for the Timer group. The top-right shows the tap-mounted 'SaniTimer' device, which begins a 30 s countdown whenever the tap is turned on. The poster encourages participants to use the timer to make sure they wash their hands for 20 s and use proper handwashing technique.

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press



Figure 3. Posters for the Precommitment group. The upper poster was placed in view of participants' workstations, so they could see it while preparing food. The lower poster was placed in view of the sink so it could be seen while participants washed their hands.

part, participants were given a very general description of the study's purpose (understanding how food handlers complete a guided food preparation task, under time constraints). A post-participation survey included a question asking participants what they thought the purpose of the study was, to check whether they had discerned the study objective. At the end of their session, all participants were fully debriefed and explicitly asked if they would like to withdraw from the study.

Procedure

The target population for this trial was UK food handlers employed at FBOs. To access our target population, a recruitment agency was used to recruit and allocate participants to testing slots, which had been pre-assigned to treatment groups. Participants were paid £60-70 for taking part. (Early morning slots offered a higher incentive as sign-up for these sessions was lower.)

Before attending the lab session, participants were emailed a link to an online pre-trial training session and post-training quiz. The training materials covered key

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

aspects of the FSA's guidance on handwashing behaviour (see Supplementary Appendix A1 for details). The post-training quiz gave feedback on any incorrect answers and participants had to input the correct answer to proceed to the next question and complete the training. We recorded their first answers for use in the analysis. All participants were required to complete their training before their lab session date, and any who did not were excluded from the study. The recruitment agency sent regular reminders to participants to complete their training.

Upon arrival at the lab, participants were directed to their assigned workstation, where they were asked to read through a study information sheet and sign a consent form.

The food preparation task was described to participants as 'the burger task'. Participants were told that they would play the part of a worker in a hamburger restaurant preparing orders from customers. The task required participants to input a fictitious order on a tablet device we provided and then prepare the burger patty and assemble the garnish. Participants were instructed to complete the burger task twice within a fixed 25 min time limit, and that they should try and prepare the burger faster the second time. A written copy of the recipe was provided to all participants as a guide (see Supplementary Appendix A2). The task was designed with a total of 8 distinct handwash points, according to FSA guidance (also shown in Supplementary Appendix A2). Cameras focused on the sinks captured footage of participants' handwashing behaviour. An additional camera was also present in each workstation as a decoy, focused on the food preparation area, to mitigate experimenter demand effects.

After completing the food preparation task and cleaning their workstation, participants completed a post-participation survey on their tablet device. This survey included questions about the participants' understanding of the experiment and some additional background information (see Supplementary Appendix A3). After the post-participation survey, participants were fully debriefed as to the purpose of the study.

Outcome measures

The primary outcome measures of the study are the frequency, duration and quality of handwash attempts made by each participant during the food preparation task.

- Duration: The mean duration of handwashing attempts where soap was used, from first application of soap to starting to rinse. (This definition necessarily means duration measures only covered handwashing instances in which the participants used soap. This outcome was therefore missing for any participants who never used soap.)
- Frequency: A count of handwashing attempts made by each participant.
- Quality:
  - A count of handwashing attempts in which soap was used.
  - A count of handwashing attempts in which the participant washed the backs of both of their hands (back-hands count).

Cameras placed above the sinks in each workstation recorded participants' handwashing behaviour in the session. The outcome measures were then derived from

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

the coded video recordings by one primary coder and a dummy coder (see Supplementary Appendix A5).

Coders were not informed of the study's hypotheses or treatments and were blind to treatment as far as was possible. The sole exception to this was the tap-mounted timer, which was necessarily visible in footage of the sink, although only from the rear so it was not clear what the device was. All coders completed training against pilot footage (which was not used in the analysis) before coding any footage from the main study. As a quality check, we randomly selected 10% of participants to be double coded and analysed inter-coder reliability on this subset.

In some instances, it was unclear to the coders whether a sequence of footage contained one or multiple distinct handwashing attempts. This occurred when participants applied soap and rinsed multiple times. We amended the code frame to include a flag identifying these instances as 'ambiguous' and ran a sensitivity analysis excluding them for each primary outcome model.

The post-participation survey probed how participants engaged with the interventions. We asked participants how long they tried to spend washing their hands each time they did so, and how they kept track of time. We also tested whether participants in the *Precommitment* group could recall the statements they had committed to. The real precommitment statements were presented in lists alongside dummy statements and participants were asked to say which they had committed to. We presented the statements across two questions – one for statements prompted at workstations and one for statements prompted at the sink (see Fig. 3).

The post-participation survey also collected background information to inform our analysis, including what participants thought the purpose of the experiment was, information about their job and workplace, and demographic information. For a full list of questions asked, see Supplementary Appendix A3.

Power/sample size

Pre-trial power analysis identified a target sample size of 207 participants ($n = 69$ in each group). This would allow us sufficient power (0.8) to detect a medium-sized effect ($\alpha = 0.05$), accounting for the need to control the family-wise Type 1 error rate using Bonferroni adjustments when comparing all arms against one another.

Statistical methods

We preregistered our analysis plan on the Open Science Framework before data collection began (https://osf.io/zpvnx).

We analysed each of the primary outcome measures listed above separately. For each, we carried out a generalised linear model (GLM) with the primary outcome as the dependent variable. For the three 'count' outcomes, we used a Poisson log-link GLM. For average duration, no transformation was necessary, so we used an identity link function.

For each outcome measure, the model includes Experimental arm as the primary independent variable of interest. In the models, experimental arm is dummy coded into two binary variables to indicate the presence/absence of the

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

two intervention sets. The primary comparisons of interest were the effect of each of the two interventions against the control arm, as set out in the hypotheses. We thus used a Bonferroni-adjusted α of 0.025 to account for the family-wise Type-1 error rate for these comparisons. The results section reports unadjusted p-values.

The two quality measures (count of handwashing occasions in which soap was used and count of handwashing occasions in which participants washed the backs of both of their hands) also include the count of handwashing attempts made as an essential covariate, to account for variation in handwashing frequency.

In addition, the initial model for each outcome measure features the following non-essential independent variables. (In this list, superscript a indicates covariates that were not listed in the initial analysis plan as counterbalancing between the treatment groups should have accounted for any systematic effects. We decided to include these additional covariates early in fieldwork in response to a higher drop-out rate than anticipated, which could have undermined our counterbalancing. As a robustness check, any best-fit models including these covariates were also checked without them.):

- Which of the three workstations the participant was assigned to;
- FBO employer type (collected in the post-trial questionnaire);
- Job type (collected in the post-trial questionnaire);
- Age (collected in the post-trial questionnaire);
- Gender (collected in the post-trial questionnaire);
- The participant's score on knowledge measures taken from the pre-trial survey;
- Whether the participants correctly guessed the purpose of the trial was to observe handwashing behaviour;
- Whether the participant finished the task;
- Which day of the week testing took place on;[a]
- Which time slot testing took place in.[a]

To arrive at a parsimonious model and avoid over-fitting, we iteratively dropped non-essential independent variables from the models as long as doing so improved goodness of fit according to Akaike's Information Criterion.

Results

Participants and balancing checks

The experiment ran between 7 March and 27 May 2022. In total, 214 participants took part over 21 sessions. We excluded 11 participants who had not completed their training and three who could be personally identified from their video footage (a requirement of the study's ethical approval). Technical issues led to lost footage for a further five participants. This left us a total of $n = 195$ (Timer: $n = 69$; Precommitment: $n = 59$; Control: $n = 67$) participants, which was below our target sample size.

Participants were aged between 17 and 64 years old ($M = 35.2$, $SD = 11.9$; excluding two participants who refused to answer this question). Eighty-three identified as male (43%) and 107 as female (55%). Four selected 'I identify in a different way' and

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

one refused to answer the question. Baseline demographic characteristics for each treatment group can be found in Supplementary Appendix B1.

Participants scored an average of 5.79 ($SD = 1.20$) out of 8 on the pre-trial training quiz. One-way ANOVA found no baseline differences in training score between treatment groups (*Timer:* $M = 5.68$, $SD = 1.27$; *Precommitment:* $M = 5.83$, $SD = 1.22$; *Control:* $M = 5.87$ $SD = 1.11$; $F(2, 192) = 0.45$, $p = 0.639$).

Eight participants (4.1%) ran out of time and were unable to finish the food preparation task fully. At the end of each session, but before debriefing, we asked participants what they thought the purpose of the experiment was. In total, 41 (21%) correctly guessed that the purpose of the experiment was to examine handwashing behaviour. The proportion of participants who correctly guessed the experiment purpose did not differ across treatment arms ($X^2(2,194) = 3.32$, $p = 0.190$).

Inter-coder reliability

To determine agreement between the two coders on the subsample selected for double coding, we ran Kendall's W for each of the four primary outcomes. The results show statistically significant agreement between the two coders on every outcome measure (handwashing attempt count: $W = 0.976$, $p = 0.007$; mean duration: $W = 0.926$, $p = 0.012$; count of attempts in which soap was used: $W = 0.998$, $p = 0.005$; count of attempts in which soap was used and the backs of both hands were washed: $W = 0.960$, $p = 0.008$).

Survey measures on engagement with the interventions

The post-trial survey probed how well participants engaged with the interventions (see Tables B2.4 to B2.9 in Supplementary Appendix B2).

Thirty-two participants in the *Timer* group (46%) said they used the countdown timer attached to the sink to keep track of how long they spent washing their hands. Of those, 28 said the timer was the *main thing* they used to keep track. The rest of the group either said they did not use the timer ($n = 31$) or could not say how long they tried to wash their hands for ($n = 6$). The most common reason for *not* using the timer was '*I didn't notice it*' ($n = 15$).

Eight participants in the *Precommitment* group (14%) correctly recalled all the statements they had committed to as part of the precommitment intervention. Forty-four (75%) correctly recalled both statements pertaining to handwashing *frequency* ('If I touch raw meat, then I will wash my hands afterwards' and 'If I am going to touch or handle any food, then I will wash my hands first') but only nine (15%) correctly recalled all three statements pertaining to handwashing *quality* ('Using soap and wetting my hands under warm running water', 'Washing the backs of my hands by rubbing against the other hand's palm' and 'Washing my fingertips by rubbing against my other hand's palm').

Duration of handwashing attempts

Descriptive statistics

Participants in the *Control* group washed their hands for 8.65 s on average ($SD = 4.76$) compared with 10.51 s ($SD = 5.94$) in the *Timer* group and 11.10 s ($SD = 6.41$) in the

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Precommitment group (Table 1). The mean duration was therefore lower than the 20 s advised by the FSA in all groups.

Confirmatory inferential analysis

Our best-fit linear regression model found statistically significant increases in duration for both interventions when compared against the *Control* group (after making a Bonferroni correction to account for the number of pairwise comparisons). The model's coefficients show the size of the increase in duration relative to the control, after accounting for other covariates (Table 2). The timer intervention was associated with a 2.20 s increase in handwashing duration ($p = 0.021$) and the precommitment intervention was associated with a 2.30 s increase in handwashing duration ($p = 0.022$).

Sensitivity analysis

Sensitivity analysis (see Supplementary Appendix B3) shows the effect of the timer intervention on duration is not robust to the exclusion of 'ambiguous' handwash attempts. (The coders were instructed to label a handwashing attempt as 'ambiguous' if it was unclear from the footage as to whether one or multiple handwashing attempts had been made. These ambiguous attempts have the potential to distort the study's primary outcomes if coded imperfectly, so re-running the primary analysis without them serves as a check of our findings' robustness.)

Post hoc *exploratory analysis*

If we exclude *Timer* group participants who said they did not use the timer from our best-fit model, the duration increase associated with the timer is larger (2.77 s, $p = 0.016$) and robust to exclusion of 'ambiguous' handwashing attempts.

Our results are broadly in line with our hypotheses: both interventions appeared to increase handwashing duration in our preregistered analysis. However, *post hoc* analyses suggest the effect of the timer intervention may be driven by a sub-group of *Timer* group participants who say they actually used the timer.

Frequency of handwashing

Descriptive statistics

Participants in the *Control* group washed their hands 6.13 times on average ($SD = 2.65$) compared with 5.48 times ($SD = 2.13$) in the *Timer* group and 6.37 times ($SD = 2.62$) in the *Precommitment* group (Table 1). The food preparation task included an expected eight handwashing occasions; only 17 participants (9%) washed their hands at least this often (*Timer group*: 6%, *Precommitment group:* 12%, *Control group:* 9%).

Confirmatory inferential analysis

Our best-fit Poisson log-link regression model found no effect of either intervention on the number of handwashing attempts made (Table 3). This is consistent with Hypothesis 1 (no effect of the timer intervention), but not Hypothesis 2 (more handwashing with the precommitment intervention).

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Table 1. Summary of handwashing primary outcome variables

	Timer ($n = 69$)	Precommitment ($n = 59$)	Control ($n = 67$)
	Mean (SD)	Mean (SD)	Mean (SD)
Mean duration	10.51 (5.94)	11.10 (6.41)	8.65 (4.76)
Handwash attempt count	5.48 (2.13)	6.37 (2.62)	6.13 (2.65)
Soap-use count	4.68 (2.23)	5.15 (2.52)	5.24 (2.44)
% Soap-use[a]	85%	80%	84%
Back-hands count	3.78 (2.37)	4.34 (2.35)	3.86 (2.50)
% Back-hands[b]	67%	65%	59%

[a]%Soap-use is the proportion of handwash attempts where soap was used (soap-use count/handwash attempt count).
[b]%Back-hands is the proportion of handwash attempts where backs of both hands were washed (back-hands count/handwash attempt count).

Table 2. GLM model with mean duration as the outcome variable

	Mean duration		
Predictors	β	95% CI	*P*
Intercept	5.94	1.38-10.50	0.011
Timer	2.20	0.34-4.06	0.021
Precommitment	2.30	0.33-4.27	0.022
Finished task: No	4.46	0.31-8.60	0.035
Purpose correct: No	−1.23	−3.16-0.69	0.210
Training score	0.68	0.02-1.34	0.042
Table (baseline = C)			
A	0.24	−1.68-2.16	0.808
B	−1.84	−3.79-0.11	0.064

Note: $n = 188$ (seven cases dropped due to missing data for one or more covariate(s)).

Post hoc exploratory analysis

Our best-fit model did find a statistically significant effect of gender: women washed their hands more often than men on average, holding relevant covariates constant. This is consistent with previous findings (Judah *et al.*, 2009), so we ran a *post hoc* model with an interaction term between gender and treatment group. The interaction model was a worse fit than the main effects model reported in Table 3 and the interaction term was not statistically significant.

Soap use

Descriptive statistics

Participants in the *Control* group washed their hands with soap 5.24 times on average ($SD = 2.44$) compared with 4.68 times ($SD = 2.23$) in the *Timer* group and

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Table 3. GLM Poisson model with handwash frequency as the outcome variable

	Handwashing attempt count		
Predictors	Exp(β)	95% CI	*P*
Intercept	5.54	3.96-7.76	<0.001
Timer	0.90	0.78-1.04	0.154
Precommitment	1.02	0.89-1.18	0.729
Male: Yes	0.88	0.78-0.99	0.033
Purpose Correct: No	0.91	0.80-1.05	0.204
Training score	1.04	0.99-1.09	0.106

Note: $n = 189$ (six cases dropped due to missing data for one or more covariate(s)).
Exp(β) is the exponentiated beta coefficient. Here, >1 indicates a positive rate of change on the outcome variable and <1 indicates a negative rate of change.

5.15 times ($SD = 2.52$) in the *Precommitment* group (Table 1). Over 80% of all handwashing attempts made during the task included soap use (*Timer group*: 85%, *Precommitment group:* 80%, *Control group:* 84%).

Confirmatory inferential analysis

Our best-fit Poisson log-link regression model found no effect of either intervention on the number of handwashing attempts using soap, controlling for the total number of handwashing attempts (Table 4). This was contrary to our hypotheses, which posited that both interventions would increase soap use.

Sensitivity analysis

The result did not change when we ran an alternative model specification as a robustness check. The soap-use model suffered from under-dispersion, so the preregistered Poisson distribution may have been a poor fit. The alternative logistic regression model treated soap use as a binary dependent variable (1 = the participant used soap every time they washed their hands, 0 = the participants washed their hands without soap at least once). We used the same approach to covariate selection as the main analysis. For no combination of covariates trialled did the model explain our data significantly better than the null (intercept-only) model unless we dropped treatment group as a factor.

Handwashing technique

Descriptive statistics

Participants in the *Control* group washed the backs of both hands 3.86 times on average ($SD = 2.50$) compared with 3.78 times ($SD = 2.37$) in the *Timer* group and 4.34 times ($SD = 2.35$) in the *Precommitment* group (Table 1).

Confirmatory inferential analysis

Our best-fit Poisson log-link regression model found no effect of either intervention on the number of handwashing attempts in which participants washed the backs of

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Table 4. GLM Poisson model with soap count as the outcome variable

	Soap use count		
Predictors	Exp(β)	95% CI	*P*
Intercept	2.01	1.47-2.76	<0.001
Timer	1.04	0.89-1.21	0.659
Precommitment	0.93	0.79-1.09	0.348
Age	1.00	0.99-1.00	0.488
Attempt count	1.18	1.15-1.21	<0.001
Purpose correct: No	0.89	0.77-1.04	0.130

Note: $n = 191$ (four cases dropped due to missing data for one or more covariate(s)).
Exp(β) is the exponentiated beta coefficient. Here, >1 indicates a positive rate of change on the outcome variable and <1 indicates a negative rate of change.

both hands, controlling for the total number of handwashing attempts (Table 5). This was contrary to our hypotheses, which posited that both interventions would improve compliance with proper handwashing technique.

Taking together the results for soap use and handwashing technique, we find no evidence of an effect on handwashing quality for either intervention treatment. This result is contrary to both Hypothesis 1 and Hypothesis 2.

Post hoc exploratory analysis

Nonetheless, over the whole sample, participants who washed their hands for longer on average *were* more likely to wash the backs of both hands ($\rho = 0.42$, $p < 0.001$). When we split the sample by treatment group, this finding holds for the *Timer* ($\rho = 0.55$, $p = <0.001$) and *Control* groups ($\rho = 0.49$, $p = <0.001$), but not in the *Precommitment* group ($\rho = 0.23$, $p = 0.111$).

Discussion

We found evidence that both a timer and a precommitment intervention led to an increase in the amount of time participants spent rubbing their hands with soap before rinsing. Participants who had access to a tap-mounted timer washed their hands for 1.9 s longer on average than *Control* group participants. Likewise, participants who committed in advance to five statements of good hand hygiene washed their hands for 2.5 s longer on average than those in the *Control* group.

We did not find any evidence of a concomitant effect on either the frequency or quality of handwashing. Participants washed their hands on average 5.97 times across the food preparation task, washed their hands with soap 5.02 times, and washed the backs of both hands with soap 3.97 times.

These results provide mixed evidence for our initial hypotheses. Hypothesis 1 predicted the observed effect of the timer on duration and the lack of effect on frequency. However, it also stated there would be an improvement in the quality of handwashing, which was not observed. Hypothesis 2 correctly predicted the increase in

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Table 5. GLM Poisson model with back-hands count as the outcome variable

	Back-hands count		
Predictors	Exp(β)	95% CI	*P*
Intercept	1.42	1.04-1.93	0.027
Timer	1.15	0.96-1.38	0.126
Precommitment	1.02	0.85-1.23	0.804
Male: Yes	1.05	0.90-1.23	0.548
Attempt count	1.20	1.16-1.24	<0.001
Purpose correct: No	0.78	0.67-0.92	0.004
Table (baseline = C)			
A	0.96	0.81-1.14	0.627
B	0.82	0.69-0.99	0.038

Note: $n = 184$ (nine cases dropped due to missing data for one or more covariate(s)).
Exp(β) is the exponentiated beta coefficient. Here, >1 indicates a positive rate of change on the outcome variable and <1 indicates a negative rate of change.

handwashing duration in the *Precommitment* group, but we also expected to see increases in handwashing frequency and quality, neither of which were observed.

Effectiveness of the timer intervention

The tap-mounted timer had the expected effect on handwashing duration, but there were two surprising findings:

Firstly, the magnitude of the effect was smaller than has been previously reported in other studies (Yu *et al.*, 2018; Her *et al.*, 2019). Yu and colleagues reported that attaching a musical timer to soap dispensers led to a 6.4 s increase in handwashing duration relative to pre-test baseline. Her and colleagues found a 4.1 s increase using the same timer device as used in this study, again relative to a pre-test baseline. Our observed mean difference (1.9 s) is smaller. Only eight participants in the *Timer* group washed their hands for the advised length of time or longer, so a ceiling effect is unlikely to explain why the timer appeared to have a smaller effect in this study than in previous studies.

There are methodological differences between this study and those previously published, which may account for some of the discrepancy in observed effect sizes. We applied a stricter definition of handwashing duration than those studies – beginning with first application of soap and ending when rinsing starts. Yu and colleagues defined handwashing as ending when participants wiped their hands and Her and colleagues defined the endpoint as when the tap was turned off. These definitional differences seem likely to underpin the higher baseline durations reported in those studies.

Another reason we might expect a smaller effect than previously reported is if our participants did not engage as well with the intervention. This would not be surprising as previously reported studies were field trials run over multiple weeks, so participants encountered the interventions repeatedly. By contrast, our participants

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

generally washed their hands 5-6 times in a 25 min task, so they had limited opportunities to familiarise themselves with the timer. Indeed, the most common reason given for not using the timer was '*I didn't notice it*'. Nearly half ($n = 31$) of *Timer* group participants said they did not use the timer to keep track of how long they washed their hands for. If we drop these participants from our main analysis, the effect of the timer becomes larger and more robust. This finding makes intuitive sense – the timer works most well for those users consciously engaging with it – but it is the result of a *post hoc* split so we cannot infer causality.

The second surprising finding was that, while the timer did appear to have the expected effect on duration, we did not see a concomitant increase in participants washing the backs of their hands. We hypothesised that the timer would increase quality for two reasons: Firstly, the poster installed by the sink in the *Timer* group showed the FSA's guidance on handwashing techniques, including explicitly prompting participants to wash the backs of their hands. Secondly, washing one's hands for longer makes it easier to follow the guidance on technique. We did observe an overall relationship between mean duration and handwashing technique, but this did not manifest as a detectable difference across treatment group. In other words, participants spent longer washing their hands, but it is not clear that they achieved better hand hygiene by doing so.

This does raise the question of whether directly targeting handwashing duration is worthwhile. Previous work using the same timer device found an effect on duration with no concomitant improvement in quality measures (Her *et al.*, 2019). Similarly, training methods focusing on duration in children led to some parts of the hands (e.g., fingertips, palms) being missed (Öncü and Vayısoğlu, 2021).

We did not observe an effect on soap use either, suggesting that the poster did not prompt participants who were not planning to use soap to do so. Soap use was generally high, being used for at least 80% of all handwashing attempts, across all groups. It should be noted that duration was only measured for handwashing attempts in which soap was used, so we did not explore the relationship between the soap use and duration.

Finally, we did not find evidence of a 'backfire' effect on handwashing frequency (Osman *et al.*, 2020). In other words, the timer did not detectably discourage handwashing by making it more onerous.

Effectiveness of the precommitment intervention

The precommitment intervention did not have the expected effects on the frequency of handwashing, the frequency of soap-use, or whether participants applied proper handwashing technique by washing the backs of their hands. It was, however, associated with an increase in mean handwashing duration. We designed the precommitment intervention to directly target frequency and quality, with any effect on duration being a secondary consequence. Our results are therefore difficult to interpret cleanly.

Firstly, we consider why the precommitment intervention did not have the expected proximate effect on handwashing frequency or quality. In general, the evidence on the effectiveness of pledges and especially written pledges is mixed, with some studies finding positive effects (Katzev and Pardini, 1987; Lokhorst *et al.*, 2013), but others finding they are not effective (e.g., Shu *et al.*, 2012).

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Perhaps participants made the commitment dishonestly and did not intend to follow through. The study offered money for taking part and committing to the prompt statements was mandatory for participants in the *Precommitment* group, so they were financially incentivised to do so. Even if the commitment were made honestly, we might not have induced a strong enough commitment for it to be effective. For instance, it has been hypothesised that people keep commitments to maintain a positive self-image (Cialdini, 2008) or reduce cognitive dissonance (Harmon-Jones and Harmon-Jones, 2012). It therefore follows that if the commitment is not strong enough to trigger self-image concerns or cause significant cognitive dissonance, it will not be effective (Sheeran *et al.*, 2005). Since signing the prompt statements was mandatory to complete the online training, participants may have clicked through without feeling strongly enough committed to trigger these mechanisms.

Even if participants did intend to follow through on their prior commitments, there may have been an 'intention-action gap' at play (Sheeran and Webb, 2016). For example, they may have underestimated how difficult adhering to their commitment was when completing the food preparation task under time pressure. Alternatively, they may have failed to remember what they had committed to. The former seems unlikely: only three participants in the *Precommitment* group said they found the task 'Somewhat difficult' and none answered 'Very difficult'. Recall failure is more plausible: while a majority of the participants in the *Precommitment* group correctly recalled each prompt statement in the post-trial quiz, only eight (14%) correctly recalled them all.

Alternatively, it has been hypothesised that people keep commitments because the commitment is worded as an 'implementation intention', which lays down an automatic process where an action is triggered in response to a situational cue (Cialdini, 2008; Lokhorst *et al.*, 2013). However, if the precommitment is not worded specifically enough (either the cue to action is not worded specifically enough or the connection between the situation and action is not specific enough) then it might not be effective (Gollwitzer *et al.*, 2010). Implementation intentions have been shown to be successful at inducing rare or one-off actions, such as voting in an election or having a flu vaccination. Our precommitments were general enough to cover all food preparation tasks. It may be that they would have worked if they were more specifically tied to our burger-making task, which raises the possibility that implementation intentions are not as effective for such general activities.

As with the timer intervention, the absence of an effect on soap-use may be at least partially explained by a ceiling effect. Soap was used for at least 80% of all handwashing attempts, across all groups.

The finding of increased duration in the absence of any detectable changes in how frequently participants washed the backs of their hands is puzzling. It is possible that another aspect of handwashing technique *did* improve and that this accounted for the change in duration. The precommitment statements and posters prompted participants to wash their fingertips, but this could not be reliably measured from test footage and so was dropped from our list of outcomes ahead of preregistration. However, this is an entirely speculative account, with no supporting evidence.

Taken together, we found no evidence suggesting that the precommitment intervention had the intended immediate effect on handwashing frequency and quality.

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

There was an increase in duration relative to control, but with no obvious causal mechanism we recommend treating this finding with caution.

Gender differences

Women washed their hands approximately 12% more often than men during the task, independent of treatment group. In general, this is consistent with the finding that gender influences handwashing rates (Kinnison *et al.*, 2004). However, we find no indication that the effectiveness of our interventions was moderated by gender as in previous hand hygiene intervention studies (Judah *et al.*, 2009; Porzig-Drummond *et al.*, 2009). No gender differences were observed for quality or duration of handwashing attempts. However, we note that we were not powered to detect interaction effects, which would have required a much larger sample size and may be an avenue for further research.

Relationship to other handwashing literatures

There is substantial literature on interventions to improve hand hygiene in Lower Middle Income Countries (LMICs) and in hospitals, but neither of these areas are very relevant to food handlers. Barriers and facilitators to hand hygiene in LMICs are very culturally specific, for example inadequate education and training, cultural beliefs, lack of resources and substandard government regulations (Delva *et al.*, 2022); and in hospitals infrastructural deficits (Ataiyero *et al.*, 2019) and overcrowding and lack of alcohol-based handrub (Loftus *et al.*, 2019). In hospitals, there are increasing efforts to supplement or replace the use of soap and water with alcohol-based hand rub, as per the WHO multimodal hand hygiene improvement strategy (Lotfinejad *et al.*, 2021).

Strengths and limitations

To our knowledge, this is the first randomised controlled laboratory trial monitoring handwashing behaviour in food handlers. This approach allowed us to control many potentially confounding sources of variation in hand hygiene to draw robust conclusions about the effectiveness of our interventions. For example, much of the existing literature compares hand hygiene before and after an intervention is introduced (Yu *et al.*, 2018; Her *et al.*, 2019). These 'pre–post' studies are often easier to implement than between-subjects randomisation, but risk confounding treatment effects with unrelated changes during the fieldwork period (e.g., prevalence of COVID-19). The laboratory setting of this study afforded us considerable control over the environment in which participants prepared food, as well as their baseline level of training. We could therefore isolate treatment effects and, taken with our secondary measures, make inferences about the mechanisms by which our interventions were operating. Finally, the emphasis on behavioural observation means our results will not suffer from the 'intention-action' gap often associated with measures of intention to act (Sheeran and Webb, 2016). By using video cameras to record handwashing, we were able to measure duration and quality of handwashing, as well as frequency. A

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

set of decoy cameras distracted participants' attention from the purpose of the trial – fairly successfully, given that only 21% correctly guessed the purpose of the experiment in the post-experiment survey.

While laboratory studies allow for more complete control of how interventions are applied, they do come with a number of critical limitations. Firstly, the study could not take place in a real professional kitchen, so the food preparation task and time pressure under which it was completed were necessarily artificial. Most obvious was the safety constraint that we could not ask participants to actually cook any food, just to prepare it. It also meant that participants were preparing food in an unfamiliar setting and in the absence of the usual social environment of their workplace. This means our findings may lack ecological validity.

The study cannot support inferences about the longevity of any observed effects. The food preparation task lasted up to 25 min, so participants in the *Timer* group did not have long to notice the timer and discern from the poster how to use it, while preparing food under a time limit. We cannot know from our results whether the effect on duration would grow stronger over time as use becomes habitual or if it would attenuate as food handlers learn to ignore it. Similarly, the training and precommitment intervention took place at most two weeks before the testing session. To determine longevity, one would need to conduct a field trial with a longer data collection period.

The nature of the population of interest means we cannot claim to have a representative sample. We opted to use a recruiter to access our sample. This had the advantage of speed and cost-effectiveness but did mean less control over the sampling process than if we had carried out all recruitment ourselves. Because our design allocated treatment at random, we would not expect sampling biases to confound our results, but it is possible that our participants were more or less sensitive to our interventions than food handlers as a population would be. The only practical alternative in the timeframe available was to recruit through an employer, which poses its own practical (will participants feel pressured to act in a given manner?) or ethical (how can we guarantee participants do not feel obliged to take part?) problems.

Finally, laboratory trials are expensive for the number of observations they yield, so (despite having a considerably larger number of units of randomisation than cluster randomised field trials) our sample size is relatively low. This means that we would not have the power to detect interactions or smaller effects of interest, or to interrogate our findings through subsample analysis. This issue was compounded by practical difficulties in recruitment. Late drop-out rates were far higher than the recruiter had anticipated, and this led us to miss our target sample size even after extending fieldwork to accommodate more testing sessions.

One major limitation of studies with small sample sizes is that they may be prone to false positive results, especially when considering the impact of interventions across multiple outcomes. A conservative solution to this problem is to adjust the threshold for statistical significance according to the number of outcomes tested, as we did when accounting for multiple comparisons between arms. In practice, however, this would reduce power so much it could preclude most laboratory trials from ever detecting a genuine effect, and even then, such adjustments may not achieve anything if the outcomes do not co-vary. Alternatively, one might drop some outcomes from the inferential analysis, but the risk of doing so is that important

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

consequences of a prospective intervention are overlooked by policymakers, including critical 'backfire' effects.

Taking together the limitations set out above, there is reason for a dose of healthy scepticism when appraising the results of any one trial. We nonetheless believe our findings have considerable value when placed in context. Firstly, we must consider the public policy context for the work. The study's primary aim was to identify interventions to improve hand hygiene in FBO workers to reduce the spread of foodborne disease. We consider the success or failure of an intervention across outcomes through that lens, and accordingly have posed the question of how useful it is to increase handwashing duration without a concomitant increase in quality.

We must also consider our results in their research context. This paper begins with a review of the available evidence, the most relevant of which is again discussed in this section. Our findings for the tap-mounted timer intervention, in particular, replicate those reported from previous studies and one might, therefore, lend them greater credence than the findings for the precommitment intervention.

Finally, we should appraise the results of public policy research on how it may inform future policy development. In the case of a laboratory trial such as this, the logical next step for a promising intervention is to carry out a field trial to gauge effectiveness in the real world. Alternatively, insights into the mechanisms by which the intervention is purported to act may lead to additional intervention development or exploratory research to identify relevant barriers and facilitators. In our view, the findings reported in this paper constitute a strong basis for further work.

Conclusion

We found evidence that the tap-mounted timer had the expected direct effect of encouraging users to spend longer washing their hands but did not have any effect on quality. This result replicates similar findings from a prior pre–post field trial using the same timer device. We also found effects on duration but not quality for the precommitment intervention, but we have not got a good explanation for this finding, which would require further research to establish whether it is a real effect and – if so – what is its cause. There were no backfire effects on frequency, increased duration did not lead to participants washing their hands less often. We are not aware of any work exploring the long-term effectiveness of the timer devices on duration; this should be addressed in future research. More fundamentally, it is unclear whether the achieved increase in duration results in any meaningful improvement in hand hygiene because of the lack of improvement in handwashing quality.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/bpp.2023.33.

Acknowledgements. We would like to thank Charles Abraham of SaniTimer for providing the tap-mounted SaniTimer® devices used in this study. We would also like to thank the Avenue Cookery School in Wandsworth, London, UK, for providing the testing venue and assisting with set-up and cleaning each day.

Funding statement. This work was supported by the Food Standards Agency (FSA), a UK government department. Employees of the FSA were involved in commissioning and conceptualising the project, and critically reviewed the paper.

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Data availability statement. The data are not publicly available due to privacy or ethical restrictions.

Competing interests. The author(s) declare none.

Ethics approval. All procedures performed in this study involving human participants were in accordance with the ethical standards of the LSE Research Ethics Committee. Ethics approval was obtained for this trial from LSE Research ethics committee (Kitchen lab study, Ref: 55135).

References

Ainsworth, B., S. Miller, J. Denison-Day, B. Stuart, J. Groot, C. Rice, J. Bostock, X.-Y. Hu, K. Morton, L. Towler, M. Moore, M. Willcox, T. Chadborn, N. Gold, R. Amlôt, P. Little and L. Yardley (2021), 'Infection control behavior at home during the COVID-19 pandemic: Observational study of a web-based behavioral intervention (Germ Defence)', *Journal of Medical Internet Research*, **23**: e22197. https://doi.org/10.2196/22197.

Angelo, K. M., A. L. Nisler, A. J. Hall, L. G. Brown and L. H. Gould (2017), 'Epidemiology of restaurant-associated foodborne disease outbreaks, United States, 1998–2013', *Epidemiology & Infection*, **145**: 523–534.

Ataiyero, Y., J. Dyson and M. Graham (2019), 'Barriers to hand hygiene practices among health care workers in sub-Saharan African countries: A narrative review', *American Journal of Infection Control*, **47**: 565–573. https://doi.org/10.1016/j.ajic.2018.09.014.

Barrett, T. and Y. Feng (2020), 'Effect of observational evaluation of food safety curricula on high school students' behavior change', *Journal of Food Protection*, **83**: 1947–1957. http://dx.doi.org/10.4315/JFP-20-086.

Birnbach, D. J., D. King, I. Vlaev, L. F. Rosen and P. D. Harvey (2013), 'Impact of environmental olfactory cues on hand hygiene behaviour in a simulated hospital environment: a randomized study', *Journal of Hospital Infection*, **85**: 79–81. https://doi.org/10.1016/j.jhin.2013.06.008.

Black, N., M. Johnston, S. Michie, J. Hartmann-Boyce, R. West, W. Viechtbauer, M. C. Eisma, C. Scott and M. de Bruin (2020), 'Behaviour change techniques associated with smoking cessation in intervention and comparator groups of randomized controlled trials: a systematic review and meta-regression', *Addiction*, **115**: 2008–2020. https://doi.org/10.1111/add.15056.

Brandstätter, V., A. Lengfelder and P. M. Gollwitzer (2001), 'Implementation intentions and efficient action initiation', *Journal of Personality and Social Psychology*, **81**: 946. https://doi.org/10.1037/0022-3514.81.5.946.

CDC (2022), *When and How to Wash Your Hands | Handwashing | CDC [WWW Document]*. CDC. https://www.cdc.gov/handwashing/when-how-handwashing.html (accessed 11.9.22).

Chapman, B., T. Eversley, K. Fillion, T. MacLaurin and D. Powell (2010), 'Assessment of food safety practices of food service food handlers (risk assessment data): Testing a communication intervention (evaluation of tools)', *Journal of Food Protection*, **73**: 1101–1107. https://doi.org/10.4315/0362-028X-73.6.1101.

Cialdini, R. (2008), *Influence: Science and Practice*, 5th edn. Boston, Mass: ed. Pearson.

Coupe, N., S. Peters, S. Rhodes and S. Cotterill (2019), 'The effect of commitment-making on weight loss and behaviour change in adults with obesity/overweight; a systematic review', *BMC Public Health*, **19**: 816. https://doi.org/10.1186/s12889-019-7185-3.

Daniel, N., N. Casadevall, P. Sun, D. Sugden and V. Aldin (2018), The Burden of Foodborne Disease in the UK 2018. 95.

Delva, S., B. Marseille, C. L. Foronda, A. Y. Solomon, T. Pfaff and D.-L. Baptiste (2022), 'Hand hygiene practices in Caribbean and Latin American countries: An integrative review', *Journal of Clinical Nursing*, https://doi.org/10.1111/jocn.16415.

do Prado, D. B., A. P. Bettoni, V. A. Correa, B. A. de Abreu Filho, L. B. Garcia, M. C. B. Tognim and C. L. Cardoso (2015), 'Practice of hand hygiene in a university dining facility', *Food Control*, **57**: 35–40. https://doi.org/10.1016/j.foodcont.2015.03.024.

Food Standards Agency (2014), Costed extension to the Second Study of Infectious Intestinal Disease in the Community: Identifying the proportion of foodborne disease in the UK and attributing foodborne disease by food commodity.

Food Standards Agency (2019), *Safer Food Better Business For Retailers*. Food Standards Agency.

Food Standards Agency (2022), *FSA Research Suggests New Higher Estimates for the Role of Food in UK Illness [WWW Document]*. Food Standards Agency. https://www.food.gov.uk/news-alerts/news/fsa-research-suggests-new-higher-estimates-for-the-role-of-food-in-uk-illness (accessed 10.21.22).

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Gollwitzer, P. M. (1999), 'Implementation intentions: Strong effects of simple plans', *American Psychologist*, **54**: 493. https://doi.org/10.1037/0003-066X.54.7.493.

Gollwitzer, P. M. and P. Sheeran (2006), 'Implementation Intentions and Goal Achievement: A Meta-analysis of Effects and Processes', in, *Advances in Experimental Social Psychology*, Academic Press, 69–119. https://doi.org/10.1016/S0065-2601(06)38002-1

Gollwitzer, P. M., F. Wieber, A. L. Myers and S. M. McCrea (2010, *How to Maximize Implementation Intention Effects, in: Then a Miracle Occurs: Focusing on Behavior in Social Psychological Theory and Research*, New York, NY, US: Oxford University Press, 137–161.

Gould, D. J., S. Creedon, A. Jeanes, N. S. Drey, J. Chudleigh and D. Moralejo (2017a), 'Impact of observing hand hygiene in practice and research: a methodological reconsideration', *Journal of Hospital Infection*, **95**: 169–174. https://doi.org/10.1016/j.jhin.2016.08.008.

Gould, D. J., D. Moralejo, N. Drey, J. H. Chudleigh and M. Taljaard (2017b), 'Interventions to improve hand hygiene compliance in patient care', *Cochrane Database of Systematic Reviews*, https://doi.org/10.1002/14651858.CD005186.pub4.

Gould, D. J., D. Moralejo, N. Drey, J. Chudleigh and M. Taljaard (2018), 'Interventions to improve hand hygiene compliance in patient care: Reflections on three systematic reviews for the Cochrane Collaboration 2007–2017', *Journal of Infection Prevention*, **19**: 108–113. https://doi.org/10.1177/1757177417751285.

Guzewich, J. and M. P. Ross (1999), *Evaluation of Risk Related to Microbiological Contamination of Ready-to-Eat Foods by Food Preparation Workers and the Effectiveness of Interventions to Minimize those Risks*. Silver Spring, MD: Food and Drug Administration, Center for Food Safety and Appli.

Haas, J. P. and E. L. Larson (2007), 'Measurement of compliance with hand hygiene', *Journal of Hospital Infection*, **66**: 6–14. https://doi.org/10.1016/j.jhin.2006.11.013.

Hansen, P. G., E. G. Larsen and C. D. Gundersen (2021), 'Reporting on one's behavior: a survey experiment on the nonvalidity of self-reported COVID-19 hygiene-relevant routine behaviors', *Behav Public Policy*, 1–18. https://doi.org/10.1017/bpp.2021.13.

Harmon-Jones, E. and C. Harmon-Jones (2012), 'Cognitive dissonance theory', *Handbook of motivation science*, **71**.

Her, E., C. Behnke and B. Almanza (2019), 'Does a water flow timer improve food handler hand washing practices in food service establishments? The effects of passive and indirect interventions', *Journal of Environmental Health*, **81**: 8–13.

Huis, A., T. van Achterberg, M. de Bruin, R. Grol, L. Schoonhoven and M. Hulscher (2012), 'A systematic review of hand hygiene improvement strategies: a behavioural approach', *Implementation Science*, **7**: 92. https://doi.org/10.1186/1748-5908-7-92.

Husain, N. R. N., W. M. W. Muda, N. I. N. Jamil, N. N. N. Hanafi and R. A. Rahman (2016), 'Effect of food safety training on food handlers' knowledge and practices: A randomized controlled trial', *British Food Journal*, **118**: 795–808. https://doi.org/10.1108/BFJ-08-2015-0294.

Inauen, J. and H.-J. Mosler (2016), 'Mechanisms of behavioural maintenance: Long-term effects of theory-based interventions to promote safe water consumption', *Psychology & Health*, **31**: 166–183. https://doi.org/10.1080/08870446.2015.1085985.

Insfran-Rivarola, A., D. Tlapa, J. Limon-Romero, Y. Baez-Lopez, M. Miranda-Ackerman, K. Arredondo-Soto and S. Ontiveros (2020), 'A systematic review and meta-analysis of the effects of food safety and hygiene training on food handlers', *Foods*, **9**: 1169. https://doi.org/10.3390/foods9091169.

Ipsos Mori (2017), Food handlers and Norovirus transmission: Social science insights.

Judah, G., R. Aunger, W.-P. Schmidt, S. Michie, S. Granger and V. Curtis (2009), 'Experimental pretesting of hand-washing interventions in a natural setting', *American Journal of Public Health*, **99**: S405–S411. https://doi.org/10.2105/AJPH.2009.164160.

Katzev, R. D. and A. U. Pardini (1987), 'The comparative effectiveness of reward and commitment approaches in motivating community recycling', *Journal of Environmental Systems*, **17**: 93–113. https://doi.org/10.2190/XV00-DD4B-EPEH-EN5R.

Kinnison, A., R. R. Cottrell and K. A. King (2004), 'Proper hand-washing techniques in public restrooms: Differences in gender, race, signage, and time of day', *American Journal of Health Education*, **35**: 86–89. https://doi.org/10.1080/19325037.2004.10603614.

Loftus, M. J., C. Guitart, E. Tartari, A. J. Stewardson, F. Amer, F. Bellissimo-Rodrigues, Y. F. Lee, S. Mehtar, B. L. Sithole and D. Pittet (2019), 'Hand hygiene in low- and middle-income countries', *International Journal of Infectious Diseases*, **86**: 25–30. https://doi.org/10.1016/j.ijid.2019.06.002.

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press

Lokhorst, A. M., C. Werner, H. Staats, E. van Dijk and J. L. Gale (2013), 'Commitment and behavior change: A meta-analysis and critical review of commitment-making strategies in environmental research', *Environment and Behavior*, **45**: 3–34. https://doi.org/10.1177/0013916511411477.

Lotfinejad, N., A. Peters, E. Tartari, C. Fankhauser-Rodriguez, D. Pires and D. Pittet (2021), 'Hand hygiene in health care: 20 years of ongoing advances and perspectives', *The Lancet Infectious Diseases*, **21**: e209–e221. https://doi.org/10.1016/S1473-3099(21)00383-2.

McFarland, P., A. Checinska Sielaff, B. Rasco and S. Smith (2019), 'Efficacy of food safety training in commercial food service', *Journal of Food Science*, **84**: 1239–1246. https://doi.org/10.1111/1750-3841.14628.

Michie, S., R. West and R. Amlôt (2020), 'Behavioural strategies for reducing covid-19 transmission in the general population', *BMJ Opinion*, **3**.

NHS (2022), *How to Wash Your Hands [WWW Document]*. NHS website. https://www.nhs.uk/live-well/best-way-to-wash-your-hands/ (accessed 11.9.22).

Olsen, S. J., L. C. MacKinon, J. S. Goulding, N. H. Bean and L. Slutsker (2000), Surveillance for foodborne-disease outbreaks, United States, 1993-1997 [WWW Document]. https://stacks.cdc.gov/view/cdc/6764 (accessed 11.24.21).

Öncü, E. and S. K. Vayısoğlu (2021), 'Duration or technique to improve the effectiveness of children' hand hygiene: A randomized controlled trial', *American Journal of Infection Control*, **49**: 1395–1401. https://doi.org/10.1016/j.ajic.2021.03.012.

Osman, M., S. McLachlan, N. Fenton, M. Neil, R. Löfstedt and B. Meder (2020), 'Learning from behavioural changes that fail', *Trends in Cognitive Sciences*, **24**: 969–980. https://doi.org/10.1016/j.tics.2020.09.009.

Porzig-Drummond, R., R. Stevenson, T. Case and M. Oaten (2009), 'Can the emotion of disgust be harnessed to promote hand hygiene? Experimental and field-based tests', *Social Science & Medicine*, **68**: 1006–1012. https://doi.org/10.1016/j.socscimed.2009.01.013.

Sheeran, P. and T. L. Webb (2016), 'The intention–behavior gap', *Social and Personality Psychology Compass*, **10**: 503–518. https://doi.org/10.1111/spc3.12265.

Sheeran, P., T. L. Webb and P. M. Gollwitzer (2005), 'The interplay between goal intentions and implementation intentions', *Pers Soc Psychol Bull*, **31**: 87–98. https://doi.org/10.1177/0146167204271308.

Shu, L., N. Mazar, F. Gino, D. Ariely and M. H. Bazerman (2012), 'Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end', *Proceedings of the National Academy of Sciences*, **109**: 15197–200. https://doi.org/10.1073/pnas.1209746109.

Staniford, L. J. and K. A. Schmidtke (2020), 'A systematic review of hand-hygiene and environmental-disinfection interventions in settings with children', *BMC Public Health*, **20**: 195. https://doi.org/10.1186/s12889-020-8301-0.

Štefančič, V. and M. Jevšnik (2020), 'Nudge tools for improving hygiene behavior among food handlers: Case study', *Journal of Food Safety*, **40**: e12836. https://doi.org/10.1111/jfs.12836.

Yu, H., J. Neal, M. Dawson and J. M. Madera (2018), 'Implementation of behavior-based training can improve food service employees' handwashing frequencies, duration, and effectiveness', *Cornell Hospitality Quarterly*, **59**: 70–77. https://doi.org/10.1177/1938965517704370.

Cite this article: Thom J *et al* (2026). The effect of timers and precommitments on handwashing: a randomised controlled trial in a kitchen laboratory. *Behavioural Public Policy* **10**(2): 292–316. https://doi.org/10.1017/ bpp.2023.33

https://doi.org/10.1017/bpp.2023.33 Published online by Cambridge University Press