

Designing nudges for the context:  
Golden coin decals nudge workplace behavior in China

*In press at OBHDP*

Sherry Jueyu Wu\*<sup>1</sup>,  
Elizabeth Levy Paluck  
Department of Psychology  
Princeton University

Author's note.

\*Princeton University, Department of Psychology, Princeton, New Jersey 08544

Corresponding author: <sup>1</sup>Sherry Jueyu Wu

E-mail: [jueyuw@princeton.edu](mailto:jueyuw@princeton.edu)

The study protocol was approved by Princeton's Institutional Review Board. Datasets and analysis code (in R) from the current research are posted at Open Science Framework (<https://osf.io/asu8h/>). Additional analyses can be found in supplementary material.

### Abstract

We conduct a field experiment in a Chinese workplace to illustrate that designing nudges requires an understanding of the motivational structure of a behavior, which includes recognizing people's subjective understandings of their local context. Workers in six production departments of a textile factory in China were unresponsive to the factory's rules and monetary incentives to throw waste in the trash can, rather than on the floor. We designed a nudge in the form of decals depicting golden coins placed on the production floors. The coins were intended to counter workers' motivation to work without pause with their motivation to keep the golden coins uncontaminated by waste, given the shared belief that golden coins are an omen for fortune and luck. Using a stepped wedge repeated within-group design, we randomized which days the coin nudge was implemented, removed, and re-implemented in each department over a period of 5 months. We collected and coded daily pictures (7,927 total) of each production floor before, during, and after these "coin nudge" implementations and removals. Waste on the floor was significantly reduced by over 20% following the first coin nudge implementation, compared to baseline. However, the coin nudge was not effective when reimplemented, after coins were removed without justification. Removing and re-implementing the coin nudge may have shifted workers' subjective interpretation of the coins. Results support the idea that nudges are not always off-the-shelf, given that they must recognize motivations and subjective interpretations within a particular context.

Word count = 242

Keywords: nudge, behavior change, construal, waste, organizational science

## 1. Introduction

Years of research and policy experience have demonstrated that nudges—interventions that neither forbid options nor significantly change economic incentives—are effective for influencing a range of important behaviors (Thaler & Sunstein, 2008; Halpern, 2015; Benartzi et al., 2017; Allcott & Rogers, 2014; Milkman, Beshears, Choi, Laibson, & Madrian, 2011). Part of the appeal of nudges is that many of those attempted have succeeded across a range of different situations. Nudges like default options (Thaler & Benartzi, 2004) and reminders (Milkman, Beshears, Choi, Laibson, & Madrian, 2011) thus have the reputation for being transferrable, off-the-shelf interventions. These interventions are designed in light of relatively universal features of human cognition, such as loss aversion, inertia, and limited self-control and cognitive capacity (for a review, see Sunstein, 2016). The “stickiness” of default rules is such an instance – default options take into account the power of inertia, suggestion, and loss aversion, and have had strong effects in various social issues involving saving, vaccination, and organ donation (Thaler & Benartzi, 2004; Chetty, Friedman, Leth-Petersen, Nielsen, & Olsen, 2014; Sunstein, 2017a; Chapman, Li, Colcy, & Yoon, 2010).

Of course, even with defaults and other off-the-shelf nudges, context matters. First, context is important because nudges, like any other behavior change intervention, must be deployed in ways that recognize the motivational forces behind a behavior (Lewin, 1943). Nudges do not only influence on cognitive forces, like mis-estimation of risk or of time. Nudges also influence behavior when they align with people’s motivation—specifically, when individuals are already motivated to choose a particular outcome, or when they are relatively indifferent and open to guidance (Sunstein, 2017b). Nudging a person against their will is less effective, and may even be unethical (Orentlicher, 2009; Rebonato, 2012). Thus, one reason why

a particular nudge may not transfer to a new context is that it may not address the same motivational equilibrium, i.e., the balance of various motivational forces that produce a behavior (e.g. Gneezy & Rustichini, 2000; Lewin, 1943). To give a concrete example, in some American communities, anti-vaccination belief and practice is resurgent (Howard & Reiss, 2018). A nudge that reminds people of the time or place of the vaccine (Thaler & Sunstein, 2008) will not be an effective tool for getting a majority of people vaccinated in that place.

A second reason why a nudge may not transfer to certain contexts is that the nudge may not be subjectively understood, or *construed*, in the same way (Paluck & Shafir, 2017; Ross & Nisbett, 1991). To have their intended effect, nudges need to be subjectively understood by members of a community as a signal to action or non-action. This might include a person's subjective interpretation of symbols like arrows, icons depicting a campsite or danger, or electronic text featuring a greyed-out option rather than a live link. Across different contexts, these symbols might mean different things, or they might carry no meaning at all. For instance, a popular example of a nudge is the fly in the urinal (literally, a small fly picture painted onto the center of the urinal), which attempts to engage men's attention to use better aim and avoid splash on the bathroom floor (Thaler & Sunstein, 2008). Iterations of this nudge employ icons that may not be universally understood, such as a golf flag (Nudge blog, 2008). Other nudges are easily understood on the streets of London or Copenhagen, such as the use of green arrows and green shoe prints painted on the pavement to nudge people to take the stairs and to throw waste in trash cans (Hansen & Jespersen, 2013; Jespersen, 2012). Indicators like shoe prints (Keep Britain Tidy, 2015) are not as relevant for other contexts of anti-littering campaigns, where citizens are not accustomed to reminders painted on pavement (Sheely, 2013).

Third, public opinion in favor of using nudges varies from country to country, depending on the perceived intent of the nudges and the liberal democratic politics of the country (Sunstein, Reisch, & Rauber, 2018). By definition, nudges are meant to be transparent and easy to avoid. This means that if citizens disagree with the means or ends of such interventions, they may actively resist the behavior or decision toward which the nudge is nudging.

China is a country where public opinion strongly favors nudging (Sunstein et al., 2018). We encountered this first hand, when working with the CEO and board of a factory located in Suzhou, China, where we wanted to conduct behavioral science research on the topic of group motivation and productivity (Wu & Paluck, under review). In contrast to the factory leadership's lukewarm interest in our own topic, we encountered great enthusiasm about the behavioral science literature regarding nudging that was part of our general presentation, specifically the studies using nudges to reduce waste. The problem faced by their garment factory was the pileup of remainders of cloth on the factory floor, which sped up workers in the short run but ultimately slowed down productivity, and required more staff for cleaning. This problem existed despite the fact that each factory worker has their own trash can for production waste and trash, placed within reach or one or two steps away. The factory leadership created a rule that workers must clean up their working space in time, reminded workers to clean up their waste with a regular bell during that time, and also implemented a monetary penalty for workers who do not keep their working environment clean in the amount of 100 *yuan*, or \$16 per month (roughly 2% of average monthly salary). None of these interventions had a noticeable effect on reducing littering behavior. How, they asked us, could they use nudges to encourage workers to put their waste in the trash cans that were right next to their work station?

We viewed their question as a way to test ideas about how to design nudges for new organizational contexts. Why a new design, and not a straightforward implementation of an existing nudge like the shoe prints to the trash approach that was successful in the UK and Denmark (Keep Britain Tidy, 2015; Jespersen, 2012)? Given evidence suggesting that residents of China are on average in favor of nudging, we could have tried an existing nudge. However, we felt that the motivational forces behind the littering behavior of factory workers were quite different than those of pedestrians in the UK or Denmark. We observed that the factory workers were extremely motivated, as piece-rate workers, to not take seconds away from their work to throw their waste into trash cans. From field observation and interviews, the workers did not view the monetary penalty as strong or certain enough to encourage them to assume responsibility for maintaining a clean working environment. Footprints, we reasoned, are more likely to guide people who are motivated to throw trash away, or are indifferent but able to follow suggestions—not people who see a new action as harmful to their bottom line. Interestingly, throwing away trash boosts productivity in the long term by preventing a pause when cleaners come through—but workers were not appreciating this in the short-term.

Instead of a nudge that guided, we wondered if we could design a nudge to create a countervailing motivation to avoid throwing waste on the floor. In short, what if the images on the floor repelled action, rather than guided it? Put a different way, what if images on the floor changed worker's perception of the floor, from a waste-space to a space where waste was inappropriate? We drew upon the first author's knowledge of how the workers would construe, or subjectively interpret, various kinds of symbols.

We decided to place images of golden coins on the factory production floors, as a means of nudging workers away from throwing waste on the floor. Golden coins have a special

significance for Chinese people in this region, and indeed in many regions of China. They are a type of omen represented by a real-world object or a symbol (geomantic omen, or 風水 in Chinese) that represents good fortune and luck. Would this type of nudge be enough to re-direct the worker's motivation—that is, their motivation to continue working without a second's interruption—to throw away their production waste?

We designed a field experiment to test this idea. Working with the total number of (six) production departments in the factory, excluded the sewing departments that were reserved for a future experiment, we implemented a stepped wedge repeated within-subjects design. Specifically, we tracked each department over a period of 20 weeks as we first implemented the coin nudge on the production floor, next removed the coins, and finally re-implemented the coins—each time on a randomly-chosen day. The randomization was conducted separately for each of the six departments. We used this method to increase our power for the small number of departments allotted to us. Our method also allowed us to test the question of whether the nudge still changed behavior when workers were aware of its status as an intervention that could be taken away and re-implemented later. That is, can a nudge still change behavior, when workers are fully aware of its function as a behavior-change mechanism?

To our knowledge, this is the first empirical study that has tested a novel nudge that was designed with religious and cultural meaning in mind. We theorize that individuals' response to a specific nudge is not static, but a dynamic process influenced by their motivation supporting the behavior and their perception of the nudge intervention. Specifically, we had three main predictions. First, we predicted that the initial implementation of golden coin decals would decrease factory workers' littering behavior, since we expected the golden coins to be a repellant,

given their cultural and religious significance. Second, we predicted that after the removal of golden coin decals from the production floors, we would observe more waste on the production floors. Third, we predicted that a re-implementation of the coin nudge intervention would again decrease the amount of waste on the production floor. In addition, we were interested to test the differences, if any, between behavior change in response to the first and second coin nudge implementations. A potential difference in the effectiveness of the first and second waves of the golden coin decals might reflect a change in individual workers' subjective perception of the coin nudges. To augment our understanding of this difference, we collected qualitative evidence using observation and interviews. We pre-registered our design and our predictions at the Open Science Framework (<https://osf.io/asu8h/>).

## **2. Method**

### *2.1. Experimental design*

We used an experimental design that tracks the behavior of interest—waste on the floor, measured daily as described below—for departments across time as they moved from an untreated to a treated experimental status, and again from an untreated to treated experimental status. The study ran for 20 consecutive weeks, from the beginning of August to the end of December in 2015. We measured a baseline untreated period in each department prior to the first coin nudge implementation. The day of that implementation, as well as the day of removal, was assigned at random independently for each department, as detailed in Table 1. We re-implemented the coin nudge simultaneously in all departments at a pre-determined date.

Specifically, decals were first implemented on each department's production floor on randomly chosen Mondays during six consecutive weeks (Monday, August 31 to Monday, September 28), and removed at randomly chosen Mondays during another six consecutive weeks



(Monday, October 12 to Monday, November 16). Therefore the duration of the first coin nudge intervention randomly varied from three weeks to ten weeks across departments. We re-implemented the nudges in all six production departments on the same date (Monday, December 7) and finished data collection three weeks later. We measured waste levels daily, using multiple photographs of each production floor from the same pre-determined vantage point. The photos were taken by Chinese research assistants and were subsequently independently coded by U.S.-based research assistants.

This experimental strategy combines features of within-subjects and between-subjects designs: at a given time, departments are compared across experimental conditions; over time, each department contributes multiple observations as a series of outcomes. Experimental designs like this play a scientific and diplomatic role in field experimentation in that they ensure treatment is not withheld from certain experimental sites (often desired by the field stakeholders) and they enable researchers to estimate average treatment effects over time, as random assignment determines the timing of each treatment (Gerber & Green, 2012).

## *2.2. Experimental stimuli and procedure*

The golden coin decal design (*Figure 1*) mimicked ancient Chinese coinage commonly used in geomantic omen—a round coin with a square hole in the middle, colored golden bronze. The image was produced by a professional designer and printed onto durable, water-proof decals. The decals were 10 centimeters diameter in size, noticeable when stuck to the factory production floor.

Coin decals were placed 2 meters apart on each department's production floor at the beginning, middle, and end of each assembly line, close to workers' standing or sitting space. Each department used 145 decals on average for each nudge implementation (the number of

decals used per site ranged from 112 to 160, depending on the size of the departments). We implemented or removed the decals on pre-determined Monday mornings before workers come to work.

In discussions with the human resource department of the factory, we came to the agreement that decals should be sent to a local Buddhist temple for a blessing before their installation (this reflected in part the human resource managers' belief in the significance of the coins). Each department's supervisor (one per department), who reported to human resources, was aware of the experiment but was unaware of specific research hypotheses. Because our research assumption was that the coins would speak for themselves, the research team made no public announcement about the decals or the blessing. The supervisors knew that the decals were blessed to protect fortune and luck and were allowed to relay this to the workers if asked. We hypothesized that the coins were significant to workers even without a formal blessing, and so we did not seek to standardize this information. Not many workers approached the supervisors for this information.

### *2.3. Measurement*

Data were recorded by 6 research assistants (RAs) recruited from the office of human resources in the factory. RAs were all young female administrative staff who had relatively little direct contact with production workers and thus were unlikely to be recognized by workers as part of factory administration. Because the RAs were selected to be similar in gender and age to our factory workers, who were predominantly young women, we hoped that the data collection would be as unobtrusive as possible.

Before the start of the experiment, the RAs were trained by the first author to take pictures with their smartphones from a number of pre-determined fixed locations in their

assigned departments. To collect baseline data for littering behavior before treatments started, RAs began taking daily pictures at those pre-determined spots three weeks before the earliest treatment date for the first nudge intervention. All RAs took pictures at lunchtime, when workers were away from production floors, in order to be unobtrusive and to avoid potential daytime variations on production floor messiness.

All pictures were coded for the amount of waste on the production floors by four U.S.-based RAs, who used a 5-point scale (1 = no waste on floor to 5 = a lot of waste on floor). Counting pieces of waste was impossible, due to piles of waste. Coders were unaware of experimental hypotheses, and also unaware of the dates and departments for which the pictures were taken. Inter-rater reliabilities were high (80% to 86% pair-wise). Any rating discrepancies larger than 1 point were discussed and resolved among coders. See the supplementary materials for samples of coded pictures.

### **3. Results**

For the six departments, 7,927 pictures were taken during the 20-week experimental period. Eight picture files were corrupted and could not be accessed by the researchers, leaving a sample of 7,919 pictures in total. During the baseline period before any treatment occurred, we observed an average of 2.48 (SD = 1.23) waste levels on the floor, which translates to approximately a little to moderate amount of waste on the floor (see Table 2). *Figure 2* displays the time trends in waste, from baseline to final coin nudge implementation, averaged across all the departments and in each one.

#### *3.1. Average treatment effect of the golden coin nudges*

Before estimating the first vs. the second implementation of the coin nudge, we tested the average treatment effect of the two nudge interventions together. We coded stimuli exposure as a

dummy variable  $T$  (0 = no coins on the floor, including the baseline and removal periods; 1 = coins on floor, including the two intervention periods), and included the departmental fixed effects in our estimation model:

$$Y_i = \beta_0 + \beta_1 T_i + \gamma \text{Department}_i + \varepsilon_i, \quad (1)$$

On average, when the coin nudge was on the production floors ( $M = 2.03$ ,  $SD = 1.06$ ), we observed significantly less waste on the floors compared to the times when the nudge had not yet been implemented or had been removed from the previous intervention ( $M = 2.41$ ,  $SD = 1.20$ ;  $\beta = -0.34$ ,  $SE = 0.03$ ,  $CI = [-0.39, -0.29]$ ,  $p < 2 \times 10^{-16}$ ; see SOM Table S1, Column 1).

Because it is possible that the difference in waste amount between the treated and untreated periods is explainable by a larger trend over time unrelated to the coin nudge, we next fitted a model where we added an autoregressive term of waste amount ( $Y_{it-1}$ ), the number of days elapsed from the beginning of data collection ( $Days_{it}$ ), and the interaction between the number of days elapsed and the dummy treatment variable:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 Days_{it} + \beta_3 (T_{it} \cdot Days_{it}) + \rho Y_{it-1} + \gamma \text{Department}_i + \varepsilon_{it}, \quad (2)$$

where  $\varepsilon_{it}$  represents the unobserved disturbance term. As predicted, the coin nudge significantly reduced waste on production floor in this model as well ( $\beta = -0.26$ ,  $SE = 0.06$ ,  $CI = [-0.37, -0.15]$ ,  $p = 4.4 \times 10^{-6}$ ; see SOM Table S1, Column 2).

### 3.2. Results of the first coin nudge implementation

We next tested the effects of the first coin nudge intervention separately. Using the same analysis model with the subset of the dataset that ends the day before the start of the second coin implementation, we found that waste on production floors was significantly reduced when the coins were initially placed on the floor ( $M_{first} = 1.95$ ,  $SD = 1.01$ ) compared with waste amount in the baseline period ( $M_{base} = 2.48$ ,  $SD = 1.23$ ;  $\beta = -0.66$ ,  $SE = .03$ ,  $CI = [-0.73, -0.59]$ ,  $p < 2 \times 10^{-}$

<sup>16</sup>; see SOM Table S2, Column 1) and to when the coin nudges were removed ( $M_{removal} = 2.35$ ,  $SD = 1.17$ ;  $\beta = -0.36$ ,  $SE = 0.03$ ,  $CI = [-0.42, -0.29]$ ,  $p < 2 \times 10^{-16}$ ; see SOM Table S2, Column 3), and, thus, to both untreated periods combined ( $M = 2.41$ ,  $SD = 1.21$ ;  $\beta = -0.44$ ,  $SE = 0.03$ ,  $CI = [-0.49, -0.38]$ ,  $p < 2 \times 10^{-16}$ ; see SOM Table S2, Column 5).

Next, we fitted model 2, where we account for time trends. As predicted, the first coin nudge intervention significantly reduced waste on production floor compared with the baseline period ( $\beta = -0.31$ ,  $SE = 0.06$ ,  $CI = [-0.43, -0.19]$ ,  $p = 5.19 \times 10^{-7}$ ; see SOM Table S2, Column 2), the removal period ( $\beta = -0.60$ ,  $SE = 0.09$ ,  $CI = [-0.77, -0.42]$ ,  $p = 1.68 \times 10^{-11}$ ; see SOM Table S2, Column 4), and both untreated periods combined ( $\beta = -0.31$ ,  $SE = 0.04$ ,  $CI = [-0.40, -0.23]$ ,  $p = 8.53 \times 10^{-13}$ ; see SOM Table S2, Column 6).

The results from this series of analyses also support our second hypothesis, that waste levels would rebound after the first coin nudge was removed. On average, waste levels returned to a level that was lower, but statistically indistinguishable from, the baseline period when time trends were taken into consideration ( $\beta = -0.19$ ,  $SE = 0.19$ ,  $CI = [-0.55, 0.18]$ ,  $p = -0.32$ ; see SOM Table S3, Column 4).

### 3.3. Results of the second coin nudge implementation

We compared the amount of waste following the second implementation of the coin nudge with the amount of waste in the baseline period and the removal period. When the coins were re-implemented ( $M_{second} = 2.17$ ,  $SD = 1.13$ ), there was no statistically significant difference in waste on the floor compared with the period after they had been removed ( $M_{removal} = 2.35$ ,  $SD = 1.17$ ;  $\beta = -0.07$ ,  $CI = [-0.0038, 0.14]$ ,  $p = 0.06$ ; see SOM Table S4, Column 1). However, we observed less waste on production floors compared with the baseline period ( $M_{base} = 2.48$ ,  $SD = 1.23$ ;  $\beta = -0.36$ ,  $SE = .04$ ,  $CI = [-0.45, -0.28]$ ,  $p < 2 \times 10^{-16}$ ; see SOM Table S4, Column 3), or

the baseline and removal periods combined ( $M = 2.41$ ,  $SD = 1.21$ ,  $\beta = -0.19$ ,  $SE = 0.03$ ,  $CI = [-0.26, -0.13]$ ,  $p = 2.63 \times 10^{-8}$ ; see SOM Table S4, Column 5).

Finally, we fitted model 2, accounting for time trends. There was no difference in the amount of waste comparing the second implementation period with the removal period ( $\beta = -0.01$ ,  $CI = [-0.09, 0.12]$ ,  $p = 0.80$ ; see SOM Table S4, Column 2), the baseline period ( $\beta = -0.28$ ,  $CI = [-0.65, 0.09]$ ,  $p = 0.13$ ; see SOM Table S4, Column 4), and both periods combined ( $\beta = -0.001$ ,  $CI = [-0.08, 0.08]$ ,  $p = 0.97$ ; see SOM Table S4, Column 6). Subsetting the data into only the two treatment periods, we observe a significant difference such that the waste level was significantly higher following the second nudge than the first nudge ( $\beta = 0.45$ ,  $CI = [0.06, 0.83]$ ,  $p = 0.022$ ).

In sum, the effect of the second implementation of the coin nudge is likely to be null. The observation of less waste on floor, especially when compared with the baseline period, may be attributed to a general time trend of waste reduction, rather than a treatment effect from the second implementation of the coin nudge. Multiple factors might contribute to a potential time trend of waste reduction. First, the factory might have lighter production demand approaching the end of the year and thus workers might have less demanding work tasks and produce less waste. Second, we might have induced some demand, especially in the early stage of the baseline period, when some workers noticed our RA's data collection efforts. Thus, it is why we control for a larger trend over time in time series experimental designs.

### *3.4. Robustness checks*

We conducted a number of additional analyses, which suggest that the quantitative findings are robust. Since the duration for the first coin nudge was randomly determined for each department and might influence the amount of waste on floor, we further controlled for each

department's duration of the first coin nudge implementation (see SOM *Table S5-S6*). In addition, we contrasted different experimental periods with the baseline using the full dataset to compare the relative size and the significance of waste changes in each experimental period (see SOM *Table S7*).

### 3.5. *Qualitative findings*

To avoid experimenter demand, we did not conduct any formal interviews during the experimental period. We gathered feedback and comments through informal conversations with a small arbitrary set of workers and supervisors after the study's conclusion. We refer to aggregate field notes in order to protect the anonymity of individual respondents.

Our qualitative observations in the factory revealed two important aspects of the workers' subjective interpretations of the coin nudges. First, we heard from both workers and from factory leadership that they believed that these golden coins had particular functional value that protected their fortune and luck. One supervisor reported to our team that a worker approached her with appreciation that the coins were there, and a wish that the coins could have been shinier and "more golden" so that they could be even more effective as geomantic omen (Field Notes, September 2015). Several workers expressed the belief that the coins were there to help them make more money and should not be contaminated. Our RAs (staff from the office of Human Resources) and departmental supervisors reported a noticeable decrease in littering behavior after the first nudge intervention and attributed it to workers taking more responsibility in maintaining a clean production floor for the sake of keeping the golden coins effective.

Second, we received many questions, and even some suspicion, from workers when they saw that the golden coin nudges were removed after the first implementation. Multiple workers in each department approached their group leaders or supervisors asking about the whereabouts

of the golden coins. They expressed a belief that the geomantic omen should not be taken away once given, and wondered how the coins could be removed after having already been introduced in the departments (Field Notes, November 2015). Several workers, including the supervisors who were aware that it was part of the experiment procedure, expressed suspicion that the coins were not supposed to be in their production department, or the coins were not blessed properly at the local temple and would not be effective. The purpose and functional value of the golden coins were questioned. Indeed, in the second coin nudge implementation, we observed less interest and attention to the coin decals, compared with the first intervention when workers were happy about the presence of the coins.

#### **4. Discussion**

In summary, we observed an important change in workplace behavior—less waste on production floors—when a nudge consisting of a “golden coin” was placed on the production floor. The coin nudge was effective in repelling waste behavior when it was initially implemented on the production floors, reducing waste by an average of 20%. The shared belief that these coins represent good fortune and luck likely led workers to refrain from throwing waste on the floor, for fear of contaminating their capacity to earn more money. In other words, the images of golden coins changed the idea of the floor from a waste-space to a space where waste was inappropriate. The effect of the golden coin nudges in reducing waste did not decay over time in the first intervention period, but it did not endure after the coins were removed from the floor. Additionally, and notably, the same coin nudges were no longer effective when they were re-implemented following a time of removal. The null effect of the second nudge intervention was possibly due to workers’ changed construal of the decals.



Why did the mere exposure to golden coin images on floor led to reduced waste, when a monetary penalty had minimal influence? We point to the motivational equilibrium of the factory workers' behavior in this organization. The piece-rate factory workers were extremely motivated to maintain high production speed and to earn more money, and the monetary penalty that was introduced was not high enough or was not enforced with enough regularity to convince workers to keep their working environment clean. Prior research also suggests social interventions may on average produce more behavioral change than material rewards and other cost-benefit interventions (Kraft-Todd, Yoeli, Bhanot, & Rand, 2015). Nudges that make certain actions salient, such as green footprints or “no litter” reminders on floor, were not likely to be effective, given that workers were *motivated to avoid* using time to throw away their waste. The images of golden coins created a strong countervailing motivation—to respect a symbol of cultural and religious meaning by not defacing it with waste. From another angle, the coins tied workers' existing motivation to earn money with a motivation to keep the floor clean, so that the geomantic omen would protect fortune and luck.

An important finding is that golden coins were no longer effective when they were removed, meaning that the effect was bound to the observation of the physical coins. Additionally, the coin nudge was no longer effective when it was re-implemented after a period of removal. Such a result is not surprising to us. Our nudge intervention affected workers' behavior based on workers' construal, i.e., individuals' subjective interpretation of the stimuli (Paluck & Shafir, 2017). Golden coins had social meanings – they were perceived as special blessings that were too sacred to be taken away easily and re-introduced without any rationale. Taking the decals away from the floor may have called into question the sacred meanings of these decals in the eyes of the worker, an idea reinforced by our qualitative observations. After

having observed the decals randomly removed and re-implemented without any justification, the workers and supervisors expressed suspicion and doubt about the value of these golden coins that were initially trusted as special blessings. Therefore when these coin decals re-appeared in the second intervention, they were no longer special—contaminating them with waste on the floor was acceptable. This finding underlines the importance of construal, or subjective interpretation, in the success of nudges and choice architecture (Thaler & Sunstein, 2008; Hansen & Jespersen, 2013; Sunstein, 2017b).

We do not interpret the apparent weakened power of the re-implemented coins to mean that nudges like this are necessarily short-lived. There was little decay in the effect of the coin nudges following the first day of their initial introduction. In other words, workers were similarly responsive to the golden coins after seeing them for 20 or 40 days. Novelty of the stimuli does not seem, in this case, to be a significant determinant of the nudge's efficacy. The functional value of the golden coins did not change in people's mind as time went on after the first implementation day. What affects the effectiveness of the golden coins was the observation that they could be easily removed and re-implemented without any justifiable reason, which we suspect had significantly changed people's interpretation of the nudges and their motivation of keeping the stimuli container (the production floor) clean.

Thus, the lessons of this research echo established findings that nudges work best when they are aligned with people's motivation and are designed with the context in mind (Thaler & Sunstein, 2008; Sunstein, 2017b). However, they also add a new message to the conversation: people's responses to a nudge can change over a relatively short term due to a changed perception of the nudge stimuli. While it is important to use appropriate nudge interventions that capitalize on the motivational structure of a behavior pattern, it is also important to monitor the

dynamic changes of individual motivations and perceptions, both before and after a specific nudge is implemented and removed. One way to build hypotheses for potential dynamic change processes is to incorporate qualitative data and field observations into research design and nudge applications.

Previous researchers have tested nudges based on reminders, warnings, information disclosure, and invocation of social norms and default rules (Schultz, Nolan, Cialdini, Goldstein, Griskevicius, 2007; Chapman et al., 2010; Sunstein, 2016), but few have ventured into the cultural and religious meanings of choice architecture. The current project, while modest in size, provides a model for researchers to translate nudge interventions in different organizational contexts. It is important to test behaviorally informed nudges and learn from those tests in order to refine the nudges (Halpern, 2015; Sunstein, 2017b), without assuming one form of nudge will generalize to others. We specifically recommend, in addition to thinking about the cognitive principles that have given rise to so many successful nudges, thinking about the motivational forces behind behavioral patterns in an organization, and the subjective interpretations of the nudges proposed.

The current project cost little (with zero long-term cost), but proved more effective than the rules, reminders, and monetary penalty the organization had attempted to date. Most importantly, this study demonstrated that nudges can be adapted to different organizational and cultural contexts by thinking about motivation and construal. To that end, we hope that this research is a template for this kind of future work.

### Acknowledgements

We thank Paige Amormino, Rachel Lim, Aneesh Rai, Calvin Yu for research assistance and our anonymous reviewers for feedback. We thank the Canadian Institute for Advanced Research (CNV1001775) and Fund for Experimental Social Sciences from Princeton University for funding. The research was approved by the Princeton University Institutional Review Board, #7262.

## References

- Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, 104, 3003-3037.
- Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., Galing, S. (2017). Should Governments Invest More in Nudging? *Psychological Science*, 28, 1041-1055.
- Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Quarterly Journal of Economics*, 129, 1141-1219.
- Chapman, G. B., Li, M., Colby, H., & Yoon, H. (2010). Opting in vs opting out of influenza vaccination. *Journal of the American Medical Association*, 304, 43–44.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York, NY: W. W. Norton & Company.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29, 1-17.
- Halpern, D. (2015). *Inside the nudge unit: how small changes can make a big difference*. London: W. H. Allen.
- Hansen, P., & Jespersen, A. (2013). Nudge and the Manipulation of Choice. A Framework for the Responsible Use of Nudge Approach to Behaviour Change in Public Policy. *European Journal of Risk Regulation*, 4, 3-28.

- Howard, J., & Reiss, D. R. (2018). The anti-vaccine movement: A litany of fallacy and errors. In A. B. Kaufman & J. C. Kaufman (Ed.), *Pseudoscience: The conspiracy against science* (pp. 195-217). Cambridge: MIT Press.
- Jespersion, S. M. (2012, Feb 16). Green nudge: Nudging litter into the bin [Web log post]. Retrieved from <https://inudgeyou.com/en/green-nudge-nudging-litter-into-the-bin/>
- Keep Britain Tidy. (2015, April). *Green footprints: An experiment to nudge people towards responsible litter disposal*. Retrieved from [http://www.keepbritaintidy.org/sites/default/files/resources/KBT\\_CFSI\\_Green\\_Footprints\\_Report\\_2015.pdf](http://www.keepbritaintidy.org/sites/default/files/resources/KBT_CFSI_Green_Footprints_Report_2015.pdf)
- Kraft-Todd, G., Yoeli, E., Bhanot, S., & Rand, D. (2015). Promoting cooperation in the field. *Current Opinion in Behavioral Sciences*, 3, 96-101.
- Lewin, K. (1943). Psychological ecology. In D. Cartwright (Ed.), *Field Theory in Social Science*. London: Social Science Paperbacks.
- Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108, 10415–10420.
- Nudge blog (2008, Apr 11). The Amsterdam urinal. Retrieved from <https://nudges.wordpress.com/the-amsterdam-urinals/>
- Orentlicher, D. (2009). Presumed Consent to Organ Donation: Its Rise and Fall in the United States. *Rutgers Law Review*, 61, 295-331.

- Paluck, E. L., & Shafir, E. (2017). The Psychology of Construal in the Design of Field Experiments. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Economic Field Experiments* (Vol. 1, pp. 245–268). North-Holland.
- Rebonato, R. (2012). *Taking liberties: A critical examination of libertarian paternalism*. London: Palgrave Macmillan.
- Rogers, T., Milkman, K. L., John, L. K., & Norton, M. I. (2015). Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. *Behavioral Science & Policy*, 1, 33–41.
- Ross, L., & Nisbett, R. E. (1991) *The Person and the Situation: Perspectives of Social Psychology*. New York: McGraw-Hill.
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Giskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5), 429–434.
- Sheely, R. (2013). *Maintaining local public goods: Evidence from rural Kenya*. Harvard Kennedy School Faculty Research Working Paper Series, RWP13-051.
- Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy*, 37, 583-588.
- Sunstein, C. R. (2016). The council of psychological advisers. *Annual Review of Psychology*, 67, 713-737.
- Sunstein, C. R. (2017a). Default rules are better than active choosing (often). *Trends in Cognitive Sciences*, 21, 600-606.
- Sunstein, C. R. (2017b). Nudges that fail. *Behavioral Public Policy*, 1, 4-25.

- Sunstein, C. R., Reisch, L. A., & Rauber, J. (2018). A worldwide consensus on nudging? Not quite, but almost: Worldwide attitudes toward nudging. *Regulation & Governance*, 12(1), 3–22.
- Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112, 164-187.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New York, NY: Penguin Group.
- Wu, S., J., & Paluck, E. L. (under review). Lewin at work: Increasing productivity through group influence.



*Table 1.* Randomization schedule for the coin nudge field experiment.

Department	Baseline	First nudge implementation	Removal of nudge	First nudge duration	Second nudge implementation	Measurement completion	Second nudge duration
Department 1	Aug 5	Sep 7	Oct 12	5 weeks	Dec 7	Dec 25	3 weeks
Department 2	Aug 5	Aug 31	Nov 9	10 weeks	Dec 7	Dec 25	3 weeks
Department 3	Aug 5	Oct 5	Oct 26	3 weeks	Dec 7	Dec 25	3 weeks
Department 4	Aug 5	Sep 14	Nov 16	9 weeks	Dec 7	Dec 25	3 weeks
Department 5	Aug 5	Sep 28	Oct 19	3 weeks	Dec 7	Dec 25	3 weeks
Department 6	Aug 5	Sep 21	Nov 2	6 weeks	Dec 7	Dec 25	3 weeks

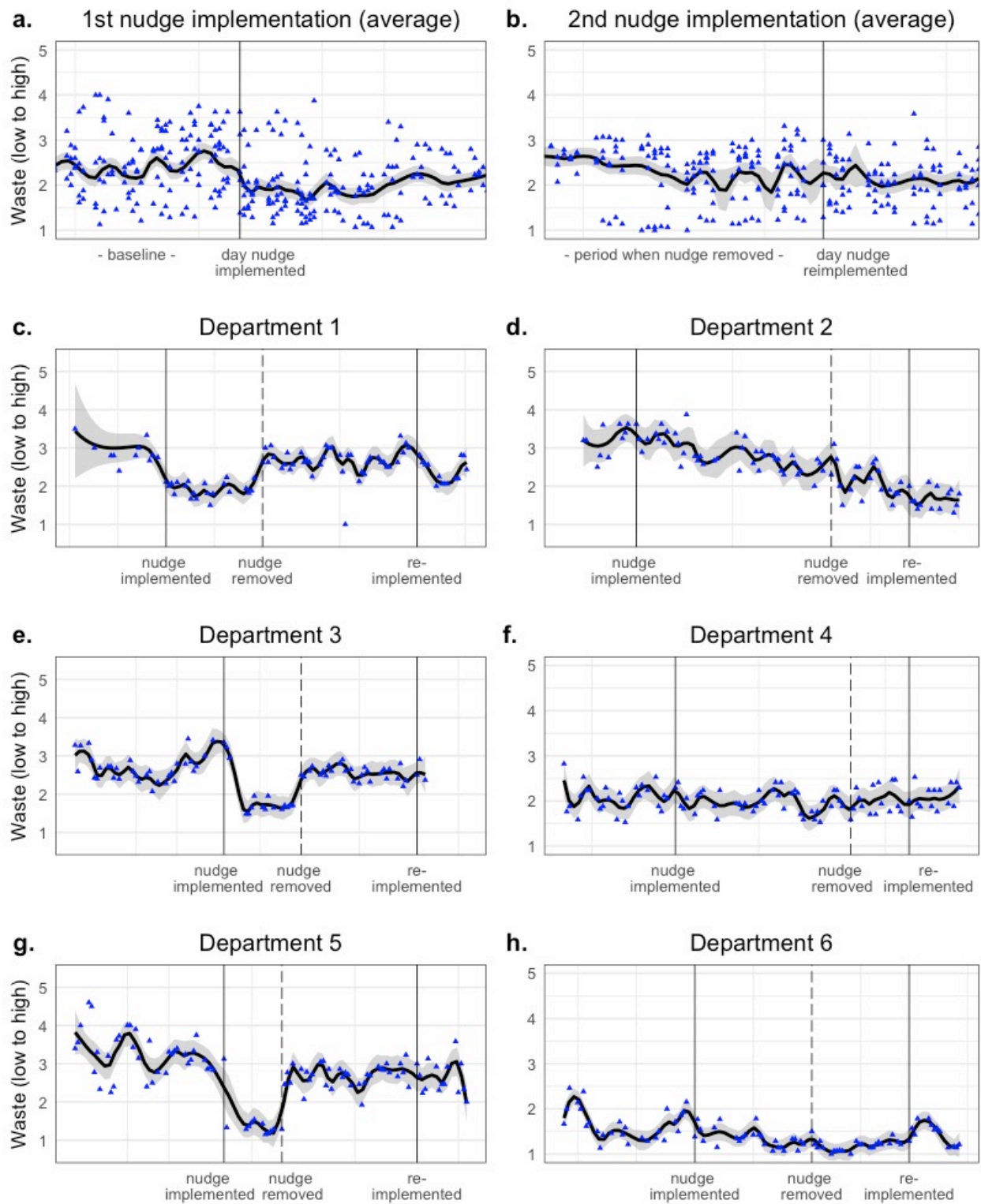
*Table 2.* Mean waste amount (1 = no waste on floor to 5 = a lot of waste on floor) for each department in different intervention periods.

Department	Baseline period	First nudge implementation	Nudge removal	Second nudge implementation
Department 1	2.88	1.92	2.69	2.33
Department 2	3.29	2.84	2.17	1.67
Department 3	2.78	1.66	2.56	2.58
Department 4	2.11	1.99	2.01	2.13
Department 5	3.30	1.43	2.66	2.73
Department 6	1.73	1.35	1.19	1.42
<b>Total</b>	<b>2.48</b>	<b>1.95</b>	<b>2.35</b>	<b>2.17</b>

*Figure 1.* The golden coin decal design, and two photographs of its implementation.



*Figure 2.* Amount of waste (y-axis) as a function of days (x-axis) during the total experimental period in each experimental department. Panels *a* and *b* depict average time trends in all six departments before and after the first and second coin nudge implementations, as a function of a standardized intervention start day. Panels *c* to *h* illustrate individual departments' daily waste amount across time. Each data point represents the average amount of waste for all the pictures taken on a single day in a department. Error bars represents 95% confidence intervals. Y-axis represents the amount of waste on the production floors on a 5-point scale as coded by independent judges (1 = no waste on floor to 5 = a lot of waste on floor).



## Supplementary Material

**Table S1. Average negative treatment effect of the coin nudges on waste.** We predicted the amount of waste on the floor with additional models that included all exposures to the coin nudge (dummy variable *Coin Nudge Exposed*: 1 = coins on floor, including the two intervention periods; 0 = no coins on the floor, including the baseline and removal periods), and each department's randomly-determined duration of the first coin nudge implementation (*Nudge Duration*; the second implementation was the same duration for all). Times when coin nudge was not present serve as the comparison. *Model 2* further controls for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *Coin Nudge Exposed*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>			
	(1)	(2)	(3)	(4)
Coin Nudge Exposed	-0.341*** (0.025)	-0.262*** (0.057)	-0.344*** (0.027)	-0.141* (0.056)
Nudge Duration			-0.021*** (0.006)	-0.002 (0.005)
Lag Waste		0.594*** (0.026)		0.778*** (0.020)
Days		-0.001* (0.0004)		0.0001 (0.0004)
Coin Nudge Exposed x Days		0.001* (0.001)		0.0005 (0.001)
Constant	2.600*** (0.032)	1.145*** (0.086)	2.529*** (0.037)	0.541*** (0.067)
Observations	7,919	7,857	7,919	7,857
Adjusted R <sup>2</sup>	0.141	0.201	0.027	0.190

**Table S2. Results for the first hypothesis—the initial implementation of the coin nudge would decrease waste.** We predicted the amount of waste on the floor with models that included the first nudge implementation (dummy variable *First Nudge*: 1 = first nudge implemented; 0 = nudge not implemented, referring to the baseline in *Models 1-2*, the removal period in *Models 3-4*, and both baseline and removal periods combined in *Models 5-6*). *Models 2, 4, 6* further control for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *First Nudge*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>					
	Baseline comparison		Removal comparison		Baseline + Removal comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
First Nudge	-0.658*** (0.035)	-0.309*** (0.062)	-0.357*** (0.033)	-0.596*** (0.088)	-0.437*** (0.029)	-0.313*** (0.044)
Lag Waste		0.563*** (0.036)		0.541*** (0.036)		0.555*** (0.029)
Days		-0.001 (0.002)		0.001 (0.001)		-0.001*** (0.0004)
First Nudge x Days		0.002 (0.002)		0.007*** (0.002)		0.005*** (0.001)
Constant	2.643*** (0.060)	1.141*** (0.110)	2.202*** (0.040)	0.993*** (0.093)	2.608*** (0.035)	1.249*** (0.090)
Observations	4,096	4,034	4,603	4,603	6,534	6,472
Adjusted R <sup>2</sup>	0.194	0.242	0.158	0.210	0.157	0.216

**Table S3. Results for the second hypothesis—removing golden coin decals from the production floors would increase waste.** We predicted the amount of waste on the floor with models that included the removal of the first nudge (dummy variable *Removal*: 1 = first nudge removed; 0 = nudge not removed, referring to the first nudge period in *Models 1-2*, the baseline in *Models 3-4*, and both baseline and first nudge periods combined in *Models 5-6*). *Models 2, 4, 6* further control for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *Removal*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>					
	First nudge comparison		Baseline comparison		Baseline + First nudge comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
Removal	0.357*** (0.033)	0.596*** (0.088)	-0.368*** (0.036)	-0.185 (0.186)	0.437*** (0.029)	0.313*** (0.044)
Lag Waste		0.541*** (0.036)		0.390*** (0.043)		0.555*** (0.029)
Days		0.001 (0.001)		0.0003 (0.002)		-0.001*** (0.0004)
Removal x Days		-0.007*** (0.002)		-0.0003 (0.003)		-0.005*** (0.001)
Constant	2.202*** (0.040)	0.993*** (0.093)	3.041*** (0.052)	1.834*** (0.152)	2.608*** (0.035)	1.249*** (0.090)
Observations	4,603	4,603	4,369	4,307	6,534	6,472
Adjusted R <sup>2</sup>	0.158	0.210	0.176	0.193	0.157	0.216



**Table S4. Results for the third hypothesis—a re-implementation of the coin nudge would again decrease the amount of waste.** We predicted the amount of waste on the floor with models that included the second nudge implementation (dummy variable *Second Nudge*: 1 = second nudge implemented; 0 = no nudge implemented, referring to the baseline in *Models 1-2*, the removal period in *Models 3-4*, and both baseline and removal periods combined in *Models 5-6*). *Models 2, 4, 6* further control for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *Second nudge*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>					
	Baseline comparison		Removal comparison		Baseline + Removal comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
Second Nudge	-0.069 (0.037)	-0.013 (0.052)	-0.362*** (0.043)	-0.281 (0.187)	-0.194*** (0.035)	-0.001 (0.042)
Lag Waste		0.284*** (0.051)		0.459*** (0.046)		0.415*** (0.037)
Days		-0.004** (0.001)		0.001 (0.002)		-0.002*** (0.0004)
Second Nudge x Days		0.003 (0.002)		-0.003 (0.002)		-0.001 (0.001)
Constant	2.533*** (0.043)	1.847*** (0.141)	2.734*** (0.066)	1.573*** (0.218)	2.660*** (0.036)	1.499*** (0.100)
Observations	3,823	3,823	3,316	3,254	5,754	5,692
Adjusted R <sup>2</sup>	0.166	0.177	0.166	0.194	0.156	0.187

**Table S5. Robustness check: Average negative treatment effect of the first coin nudge on waste.** We predicted the amount of waste on the floor with additional models that included the first nudge implementation (dummy variable *First Nudge*: 1 = first nudge implemented; 0 = nudge not implemented, referring to the baseline in *Models 1-2*, the removal period in *Models 3-4*, and both baseline and removal periods combined in *Models 5-6*), and each department's randomly determined duration of the first coin nudge implementation (*Nudge Duration*; the second nudge coin implementation was the same duration for all). *Models 2, 4, 6* further control for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *First Nudge*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>					
	Baseline comparison		Removal comparison		Baseline + Removal comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
First Nudge	-0.575*** (0.036)	-0.211*** (0.061)	-0.443*** (0.034)	-0.394*** (0.081)	-0.447*** (0.031)	-0.195*** (0.042)
Nudge Duration	0.028*** (0.007)	0.008 (0.007)	0.031*** (0.007)	0.015 (0.008)	-0.008 (0.006)	-0.003 (0.006)
Lag Waste		0.775*** (0.026)		0.765*** (0.026)		0.776*** (0.021)
Days		0.001 (0.002)		0.002 (0.001)		0.00001 (0.0004)
First Nudge x Days		0.001 (0.002)		0.005** (0.002)		0.003* (0.001)
Constant	2.320*** (0.049)	0.515*** (0.087)	1.727*** (0.056)	0.288*** (0.071)	2.453*** (0.039)	0.555*** (0.065)
Observations	4,096	4,034	4,603	4,603	6,534	6,472
Adjusted R <sup>2</sup>	0.057	0.229	0.035	0.196	0.035	0.203

**Table S6. Robustness check: Average null effect of the second coin nudge implementation on waste.** We predicted the amount of waste on the floor with additional models that included the second nudge implementation (dummy variable *Second Nudge*: 1 = second nudge implemented; 0 = no nudge implemented, referring to the baseline in *Models 1-2*, the removal period in *Models 3-4*, and both baseline and removal periods combined in *Models 5-6*), and each department's randomly determined duration of the first coin nudge implementation (*Nudge Duration*; the second nudge implementation was the same duration for all). Times when coin nudge was not present serve as the comparison. *Models 2, 4, 6* further control for the *Lag Waste* (the amount of waste in the day prior), the number of *Days* elapsed from the beginning of data collection, and the interaction between the number of *Days* elapsed and the dummy treatment variable *Second nudge*. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste</i>					
	Baseline comparison		Removal comparison		Baseline + Removal comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
Second Nudge	-0.217*** (0.042)	-0.308 (0.186)	-0.069 (0.040)	0.053 (0.052)	-0.135*** (0.037)	-0.020 (0.042)
Nudge Duration	-0.090*** (0.009)	-0.017 (0.009)	-0.091*** (0.009)	-0.022* (0.009)	-0.090*** (0.007)	-0.023** (0.007)
Lag Waste		0.731*** (0.033)		0.722*** (0.033)		0.737*** (0.025)
Days		0.003 (0.002)		-0.002 (0.001)		0.00000 (0.0004)
Second Nudge x Days		-0.005* (0.002)		0.001 (0.002)		-0.002 (0.001)
Constant	3.005*** (0.058)	1.027*** (0.197)	2.792*** (0.068)	0.774*** (0.112)	2.920*** (0.043)	0.763*** (0.084)
Observations	3,316	3,254	3,823	3,823	5,754	5,692
Adjusted R <sup>2</sup>	0.045	0.177	0.032	0.152	0.035	0.167

**Table S7. Robustness check: Waste was significantly reduced in the first nudge implementation, not in times when nudge was removed or reimplemented.** We predicted the amount of waste on the floor with additional models that included all categories of treatment periods (dummy variable *First Nudge*: 1 = first nudge implemented, 0 = otherwise; *Removal*: 1 = first nudge was removed, 0 = otherwise; *Second Nudge*: 1 = second nudge implemented, 0 = otherwise), and departmental fixed effects. The baseline serves as the comparison. *Model 2* further controls for the *Lag Waste* (the amount of waste in the day prior), and the number of *Days* elapsed from the beginning of data collection. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

	<i>Amount of waste:</i>	
	(1)	(2)
First Nudge	-0.568*** (0.035)	-0.182*** (0.047)
Removal	-0.298*** (0.034)	0.018 (0.068)
Second Nudge	-0.425*** (0.039)	-0.012 (0.089)
Lag Waste		0.580*** (0.026)
Days		-0.001 (0.001)
Fixed effects	Yes	Yes
Constant	2.815*** (0.041)	1.184*** (0.086)
Observations	7,919	7,857
Adjusted R <sup>2</sup>	0.151	0.202