Easier said than done. Do defaults and reminders affect public workers' knowledge of guidelines?

Nicola Belle, Paola Cantarelli, Chiara Barchielli & Paul Battaglio

Published online: 27 Mar 2023.

Submit your article to this journal ↗

Article views: 231

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Easier said than done. Do defaults and reminders affect public workers' knowledge of guidelines?

Nicola Belle[a], Paola Cantarelli[a], Chiara Barchielli[a], and Paul Battaglio[b]

[a]Management and Health Laboratory, Institute of Management, Scuola Superiore Sant'Anna; [b]School of Economic, Political and Policy Sciences, University of Texas at Dallas

ABSTRACT

Building on recent developments in behavioral public administration theory and methods, we conduct an online randomized controlled trial to study how defaults and reminders affect the performance of 5,303 public healthcare professionals on a test about the appropriate use of gloves. When incorrect answers are pre-populated, thus setting incorrect defaults, participants are more likely to err notwithstanding the fact that they are asked to double check the pre-populated answers. Conversely, when correct answers are pre-populated, thus setting correct defaults, subjects are less likely to err and they tend to perform better than their peers taking the non-pre-populated version of the same questions. Participants receiving either a visual or a textual reminder about the appropriate use of gloves right before the test outperform their counterparts in a control group. We also find that visual aids are more effective than textual reminders.

Introduction

Public administration and management scholarship has recently appealed for the use of insights from psychology to address pressing public human resource issues by leveraging micro-level mechanisms that drive individual attitudes and behaviors (Cantarelli et al., 2020). The ability of public organizations to fulfill their mission often depends on the day-to-day decisions and behaviors of their workers who are directly involved in public service delivery. From this perspective, boosting public professionals' knowledge of guidelines—a prerequisite of compliance—seems imperative to foster appropriateness and service quality. However, this is easier said than done. Recent work shows that, on average, it takes more than 17 years to implement guidelines in healthcare organizations and that only about half of these guidelines are widely adopted (Bauer et al., 2015).

Designing guidelines to encourage desired behaviors is common practice in public organizations across professions, including government managers and employees (Imai et al., 2013), welfare workers (Riccucci et al., 2004), teachers (Siciliano et al., 2017), and healthcare staff (Exworthy & Powell, 2004; Nagtegaal et al., 2019). This is especially true for health care professionals where guidelines geared toward directing and maintaining an appropriate use of personal protective equipment is widespread (Kang et al., 2017; World Health Organization, 2016). In fact, recent evidence shows that knowledge and compliance with such protocols enhance the safety of

CONTACT Nicola Belle ✉ nicola.belle@santannapisa.it 🖅 Management and Health Laboratory, Institute of Management, Scuola Superiore Sant'Anna, Pisa, 56127, Italy.

employees, patients, or both against the transmission of infectious pathogens (Brown et al., 2019; Verbeek et al., 2020). For example, the appropriate use of gloves is a cost-effective measure for infection control and prevention, while the inappropriate use of gloves may waste resources and has no bearing on the reduction of cross-transmission for infections. Given the serious threat posed by incorrect usage of gloves in healthcare settings, international institutions (World Health Organization, 2009:2016) and government agencies (Center for Disease Control and Prevention, 2019) sponsor instructional campaigns and behavioral interventions to encourage acceptable glove practices among healthcare professionals during service delivery in the workplace.

Building on recent developments in behavioral public administration theory and methods (Battaglio et al., 2019; Cantarelli et al., 2020; DellaVigna & Linos, 2022; Grimmelikhuijsen et al., 2017), we test two cognitive mechanisms, namely defaults and reminders, that public organizations and their managers can leverage to enhance their employees' knowledge of guidelines. We do so by experimentally manipulating defaults and reminders in the context of an online, randomized controlled trial with 5,303 public healthcare professionals taking a test about the appropriate use of gloves. Similar to behavioral research in other areas (Fonseca and Grimshaw 2017; van Dijk et al. 2020), we manipulate defaults by pre-populating information into the experimental materials. Additionally, we test and compare the impacts of both textual and visual reminders. In other words, the main objective of our study is to test the impact of defaults—in the form of correctly or incorrectly pre-filled responses—and reminders—textual or visual—on an objective test of knowledge of a regulation that healthcare professionals are obliged to know by law. In this context, both defaults and reminders are typical examples of "supposedly irrelevant factors" (Thaler 2015:9) that should not, but in fact, can significantly affect the performance of healthcare professionals.

Our work offers novel experimental findings with actionable implications for public management scholars and practitioners interested in using defaults and reminders to shape the workplace environment in which public employees make decisions (Nagtegaal et al. 2019). Notwithstanding the usual limitations that threaten the external validity of most experimental work of this ilk, our evidence scores high on the internal validity of inference. In other words, we are able to estimate the causal impact that our manipulations of defaults and reminders generate on a large sample of public healthcare professionals' knowledge of the appropriate use of gloves.

Behavioral determinants of knowledge of and compliance with guidelines

Behavioral scientists (Battaglio et al., 2019; Cantarelli et al., 2020; DellaVigna & Linos, 2022; Grimmelikhuijsen et al., 2017) and practitioners (Bhanot & Linos, 2020; Choueiki, 2016) are increasingly employing insights from psychology to address longstanding challenges in public human resource management. Defaults and reminders may be cost-effective management heuristics that leverage micro-level cognitive mechanisms to enhance compliance with policy guidelines. The global coronavirus pandemic provides us with an actionable opportunity to employ evidence from behavioral public administration; specifically, the implications of appropriate glove use among healthcare professionals. Indeed, knowledge of guidelines is often a prerequisite of compliance, thus ultimately affecting successful implementation.

Defaults

The default effect is a cognitive mechanism whereby decision makers who are asked to make a choice among alternatives tend to stick to the status quo option or to any choice that is pre-populated for them. Defaults are "the most powerful nudge we have in our arsenal" because they represent "what happens if [we] do nothing, [which] we're really good at" (Thaler, 2017:26.10). Although providing a comprehensive overview of the literature on defaults is beyond the scope of our study, it is worth noting the abundance of studies that explore the complex relationship

between defaults and active choice, whereby individuals are required to make explicit choices for themselves (Thaler & Sunstein, 2008; Carroll et al., 2009; Moseley & Stoker, 2015).

Extant research in public policy and management investigates the effect of choosing the status quo across decision areas, such as election administration (McDaniels, 1996; Moynihan & Lavertu, 2012), service provider selection (Belle et al., 2018; Jilke, 2015; Jilke et al., 2016), and healthcare policy making (Moseley & Stoker, 2015; Shpaizman, 2017; Sunstein et al., 2019; Thomann, 2018). Specific to the healthcare literature, recent work investigates the impact of defaults on physicians' decisions regarding prescription medication (Blumenthal-Barby & Krieger, 2015; Saposnik et al., 2016). For instance, a recent pre-post intervention conducted in several hospitals demonstrates that lowering prescription defaults for postoperative opioids in the electronic health records from 30 to 12 pills decreases by more than 15% the number of opioids actually prescribed (Chiu et al., 2018). Likewise, a recent pre-post study shows that the overall generic prescribing rate increases from about 75% to 98% when the generic equivalent medication is the default option in the electronic health records (Patel et al., 2016). Research examining the effect of defaults on the behaviors of other healthcare professionals, such as nurses and nurse assistants, are exceedingly rare. A recent scoping review of 124 nudges aimed at encouraging the use of evidence-based medicine among healthcare professionals found only 7% of studies focused on the impact of modifying default settings and called for further research into the determinants of compliance with hand hygiene guidelines among nurses (Nagtegaal et al., 2019).

The interdisciplinary literature on the effect of defaults has been inspired by the groundbreaking empirical work of Johnson and Goldstein (2003) and Thaler and Benartzi (2004). In their study "Do defaults save lives?," Johnson and Goldstein (2003) show that the proportion of citizens enrolled in organ donation registries is impressively higher—between 86 and 100%—in countries where all citizens are automatically enrolled unless they request cancelation (opt-out system). Conversely, the rate is much lower—between 4 and 28%—in countries where no one is enrolled unless they explicitly consent (opt-in system). Even though the decision is the same and no choice is prohibited at the individual level, the take up of a desired behavior is impressively higher under opt-out policies relative to opt-in policies. Thaler and Benartzi (2004) applied the same logic to pension plan enrollment. As "passivity is one of humans greatest skills" (Thaler, 2017:26.10), implementing a policy with automatic—as opposed to voluntary—enrollment in a savings plan significantly increased employee savings. These are classic examples of how defaults in public policy may cause striking differences (both positive and negative) in individual conduct and policy outcomes. Furthermore, these studies demonstrate the differences in behaviors generated by default options have both economic and social consequences.

Scholars have shown that the default effect not only applies to behaviors that are more or less desirable, but also extends to situations where objectively correct or objectively incorrect options exist. For example, recent work has investigated the impact of correctly or incorrectly pre-populated tax forms onto tax compliance (Fonseca and Grimshaw 2017; van Dijk et al. 2020). A randomized experiment with citizens by Fonseca and Grimshaw (2017) demonstrates that incorrect defaults in tax forms, in the form of incorrectly pre-populated fields, leads to significant drops in compliance and tax revenues. Similarly, in a laboratory experiment with students, van Dijk et al. (2020) find that correctly pre-populated tax returns promote compliance, whereas incorrectly pre-populated returns undermine it. Our study seeks to expand previous research that manipulates defaults by pre-populating either correct or incorrect information into the experimental material in the context of real public healthcare professionals taking a test on the correct use of gloves. To this end, we formulate and experimentally test the following hypotheses:

H1a: Incorrect defaults, set by pre-populating incorrect answers, decrease public healthcare professionals' performance on a test about the use of gloves.

H1b: Correct defaults, set by pre-populating correct answers, increase public healthcare professionals' performance on a test about the use of gloves.

Reminders

Reminders communicate relevant content in a timely fashion (DellaVigna & Linos, 2022). Providing reminders does not make previously unknown or inaccessible information available to decision makers. Rather, reminders make familiar content more salient by moving it into a decision maker's attention focus (Münscher et al., 2016). Public managers can purposefully employ reminders as a "direct follow-up to an initial intervention in order to reinforce its effects" (Sanders et al., 2018).

The interdisciplinary literature shows that reminders tend to be effective in directing behaviors across myriad policy areas, including, public manager and employee implementation of programs (Andersen & Hvidman, 2021), household energy conservation (Allcott & Rogers, 2014), patient dental care (Altmann & Traxler, 2014), personal savings (Karlan et al., 2016), delinquent fine collection (Haynes et al., 2013), participation in elections (Dale & Strauss, 2009; Green & Gerber, 2019), and social benefits claims (Bhargava & Manoli, 2015). In terms of healthcare, the scoping synthesis by Nagtegaal et al. (2019) found that 23% of the nudges they reviewed focused on providing reminders to healthcare professionals to improve adherence to evidence-based medicine.

Although reminders can have different formats, there is a dearth of information on the use of alternative formats in public administration and public personnel management research (Baekgaard et al., 2019; Fryer et al., 2009; Olsen, 2017). With respect to the display of data, common sense rather than evidence has traditionally guided policy (Hildon et al., 2012). This lack of a coherent strategy is problematic as the transmission of information is rarely if ever neutral. Indeed, the choice among alternative formats may predictably alter decisions (Rahman et al., 2017). In the public sector context, we know very little about the use of visualization as a heuristic for decision-making (Isett & Hicks, 2018). However, support for the enhanced efficacy of data visualization, compared to other forms of information presentation, is gaining traction among scholars (Bresciani & Eppler, 2009; John & Blume, 2018; Michie & Lester, 2005; Slingsby et al., 2014). Recently, a large-scale survey experiment with responses from 1,406 Italian local politicians demonstrated that presenting information graphically rather than textually mitigates framing biases (Baekgaard et al., 2019). Combining evidence about the effects of reminders and information formats, we formulate the following hypotheses:

> H2a: Reminders increase public healthcare professionals' performance on a test about the use of gloves.

> H2b: Visual reminders increase public healthcare professionals' performance on a test about the use of gloves more than textual reminders.

Methods

We conducted our randomized controlled trial (RCT) in the context of an anonymous online test on the appropriate use of gloves that 5,303 healthcare workers from a public regional healthcare system in Italy completed between mid-October and mid-November 2019. Participants were recruited among 7,019 public professionals who had previously completed an employee satisfaction survey. Upon completion of the employee viewpoint survey, subjects were invited to take our RCT survey by clicking on an anonymous link that would redirect them to a Qualtrics questionnaire, which was distinct and independent from the preceding survey. Participation in the experimental trial was voluntary and responses were anonymous.

Experimental task

All subjects took a test about the appropriate use of gloves in healthcare settings. Respondents were presented with seven situations and asked to select those in which the use of gloves is

appropriate according to official guidelines outlined in a specific regulation (i.e., Executive Decree 19154/2017) that is required knowledge of all public healthcare professionals in our sample. That said, respondents' previous knowledge of the decree number and/or its very content is not known in the context of this study. The seven situations include three circumstances in which the use of gloves is objectively incorrect and four in which the use of gloves is objectively correct. Therefore, incorrect and correct answers are mutually exclusive and belong to two disjoint sets. In other words, incorrectly and correctly pre-populated responses are systematically linked to different items. This could not be otherwise because Executive Decree 19154/2017 provides two mutually independent lists, one reporting circumstances in which wearing gloves is correct (e.g. "Predictable contact with blood and/or biological fluids, including injured skin and mucous membranes, excluding sweating") and the other listing circumstances in which wearing gloves is incorrect (e.g. "Handling of test tubes, bags, containers, that do not contain biological fluids and/or chemicals"). Therefore, in any of the circumstances listed in Executive Decree 19154/2017, the use of gloves is either correct or incorrect. To comprehensively evaluate the knowledge of Legislative Decree 19154/2017, the test presents a series of circumstances selected from the two lists and asks the respondents to indicate when the use of gloves is correct.

Our RCT employs a between-subjects design, whereby there are multiple conditions and different units (i.e., healthcare professionals) in each condition. More precisely, subjects were randomly allocated to one of six conditions resulting from the combination of the two factors explained below.

Factor 1: default (no, yes)

Subjects were randomly assigned to one of two levels. A random half of subjects were assigned to a control condition (i.e., *no default* or, equivalently, *no pre-populated*) and took the regular test, which did not include any pre-populated answers. In other words, participants in the *no default* group were presented with a list of seven situations, none of which had been pre-populated, and asked to select those in which they thought the use of gloves was appropriate. Figure A1 in Appendix A reports the scenario of the *no default* treatment. The quiz version presented to their counterparts in the *default* or *pre-populated* condition featured four pre-populated answers, two of which were correct and two incorrect. In addition to the same instructions presented to subjects in the *no default* arm, professionals in the *default* arm were warned to deselect any incorrectly pre-populated answers. Figure A2 in Appendix A reports the scenario of the *default* treatment. The fact that incorrectly and correctly pre-populated responses are systematically linked to different items does not invalidate the comparison of each default type—correct or incorrect—and the control group. Correctly and incorrectly pre-populated answers necessarily relate to different items because the same circumstance cannot be both true and false. In other words, this could not be otherwise in the context of our study. Our experimental manipulations of the default factor entail that subjects in both arms make an active choice. For participants in the control condition, the active choice consists in selecting the answers they think are correct. For their counterparts in the treatment arm, the active choice consists in selecting the answers they think are correct and deselect any incorrectly pre-populated answers.

Factor 2: reminder (no, textual, visual)

Subjects were randomly assigned to one of three levels. Participants in the control arm were prompted to base their responses on Executive Decree 19154/2017 and did not receive any reminders about the content of the legislation (Figure A3 in Appendix A). Therefore, professionals in the control took the test exclusively relying on any previous knowledge of guidelines on gloves use. Right before taking the test, a random third of subjects were shown for 60 seconds an excerpt from the actual legislation (i.e., Executive Decree 19154/2017) listing in a paragraph

format the circumstances in which gloves must be used in healthcare settings (Figure A4 in Appendix A). A random third of subjects were shown for 60 seconds a leaflet to be distributed in hospitals that uses visual aids to indicate in which situations gloves should and should not be used in healthcare operations based on Executive Decree 19154/2017 (Figure A5 in Appendix A). Thus, professionals randomly assigned to the treatment arms took the test after being exposed to a reminder of the content of the legislation, either in a textual or visual format. The rationale behind the selection of these operations for the textual and visual reminders lies in maximizing realism and ecological validity. In fact, the material included in the textual and visual treatments in our RCT was prepared by the regional healthcare system for field use.

As far as the levels of our reminder factor are concerned, an important caveat merits explicit consideration at this point. Precisely, if none or only few respondents know the content of the instruction, our treatment—be it textual or visual reminder—may entail providing new information. At the same time, however, we should emphasize that knowledge of the regulation is mandatory for all workers. Thus, together with the partner institution, we moved from this fairly reasonable assumption and designed all procedures "as if" all employees knew it.

Measures

We computed standardized test scores awarding +1 point each time a correct answer was flagged, −1 point each time an incorrect answer was flagged, and 0 point otherwise. Standardized scores on the full test, on the incorrect answers that might be pre-populated, on the correct answers that might be pre-populated, and on the answers that could not be pre-populated served as our four dependent variables. We use standardized scores to ease the comparison of results across research projects. Nevertheless, to enhance transparency and completeness, Table B1 of Appendix B reports the average unstandardized test scores and standard deviations by type of answers, separately for each experimental arm.

Findings

In our sample of 5,303 public healthcare employees, about 86% are nurses and 14 nursing assistants. Of the total population in our sample, approximately 75% are female, 16% male, and 9% preferred not to say. Subjects in our study are distributed as follows in terms of age groups: about 3% is younger than 30, 14% between 30 and 39, 28% between 40 and 49, 40% between 50 and 59, 6% older than 60, and 9% did not provide age information. As to organizational settings, about 52% of our respondents work in hospitals, 22% in ambulatory care, 18% in teaching hospitals, and the remaining 8% did not provide information regarding their place of work. Due to randomization, the distributions of respondents by job family, gender, age group, and organization type across the six experimental groups are statistically indistinguishable from each other.

The effects of defaults

The left most panel of Figure 1 displays the average standardized test scores on the different typologies of answers that were included in the test, along with 95% confidence intervals, separately for subjects who took the test without any pre-populated answers and for participants who answered the version of the same test that featured four pre-populated answers. As to the full test, our results show that professionals in the default condition underperformed their peers in the control group by .27 standard deviations ($p < .0005$). In other words, pre-populating two correct and two incorrect default answers, which professionals were asked to double check and possibly change, caused a sizable deterioration in test performance.
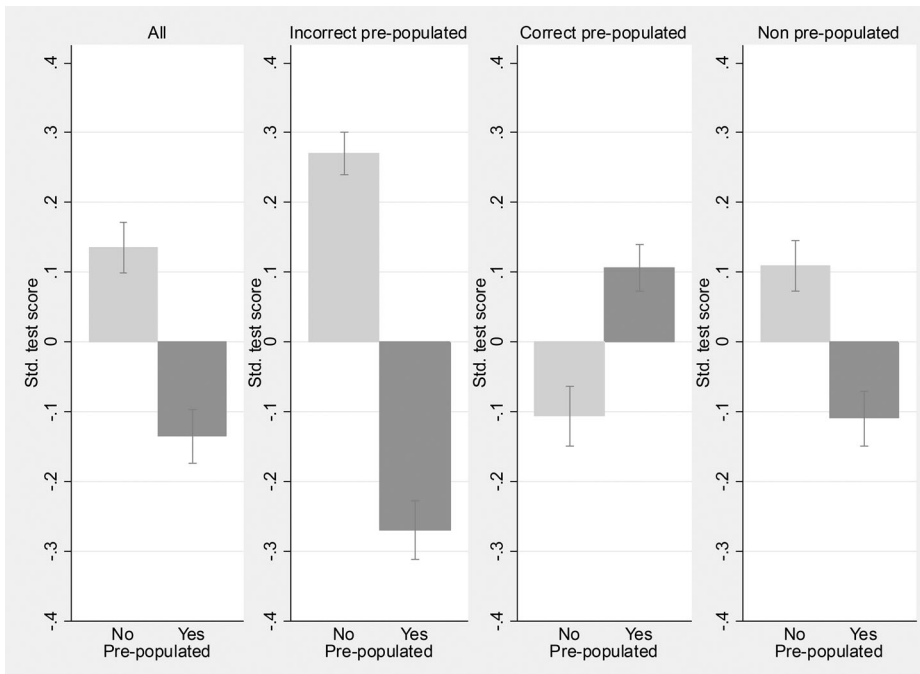
Figure 1. Standardized test scores, by typology of answers, by test version (with or without pre-populated answers).
Notes: (i) Error bars represent the 95-percent point-wise confidence interval. (ii) Estimates from the linear regression model that underlies the figure are reported in Appendix B, Table B2.

To appreciate the differences in overall test scores between subjects who took the regular test relative to their peers who took the version with pre-populated answers, we separately analyzed participant performance on the three types of answers included in the test (i.e., the two incorrect answers that might be pre-populated, the two correct answers that might be pre-populated, and the three answers that were not pre-populated). The incorrect pre-populated panel of Figure 1 shows the average standardized test scores on the two incorrect answers that were pre-populated for subjects in the default arm of our RCT but not for their counterparts in the no default group. The graph suggests that healthcare workers in our sample were more likely to err in the presence of an incorrect default. More precisely, relative to the two incorrect answers that might be pre-populated, subjects in the default arm of the experiment underperformed their counterparts in the control condition by .54 standard deviations ($p < .0005$). The results, thus, provide support for hypothesis H1a.

The correct pre-populated panel of Figure 1 mirrors the incorrect pre-populated panel and represents the average standardized test scores on the two correct answers that might be pre-populated, separately for subjects in the default and in the no default groups. As predicted by hypothesis H1b, a correct default increased professionals' performance on those two items by .21 standard deviations ($p < .0005$), though the difference between the two experimental arms is smaller relative to that portrayed in the incorrect pre-populated panel of Figure 1. In other words, the negative impact of incorrect defaults appears to outweigh the gain from correct defaults.

Then, subjects in the default arm of our RCT underperformed by .22 standard deviations ($p < .0005$) relative to their peers in the no default group on the three answers that were not pre-populated. This seems to be in line with extant scholarship showing that both pre-populated and not pre-populated options in a choice set are affected by default settings (Johnson and Goldstein 2003). A candidate explanation for this finding might include a reduced attention that subjects pay in the presence of defaults. As we explain in the Limitations section, this hypothesis cannot be tested through our experimental design and calls for future work.
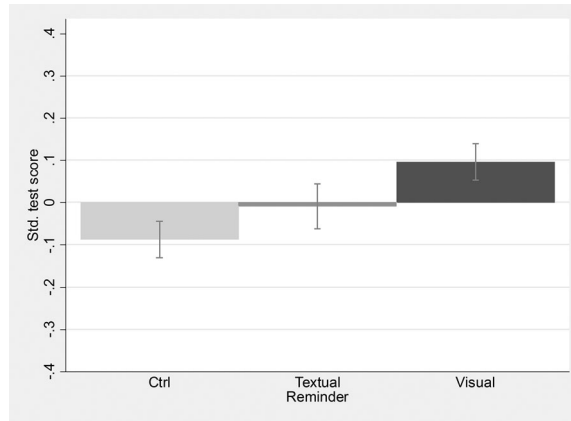
Figure 2. Standardized scores on the full test, by type of reminder.
Notes: (i) Error bars represent the 95-percent point-wise confidence interval. (ii) Estimates from the linear regression model that underlies the figure are reported in Appendix B, Table B3.

Overall, 12% of participants in the default condition did not change any answers at all. The percentage of subjects who completely stuck to the default was 95% for the correct defaults and 28% for the incorrect defaults ($p < .0001$).

The effects of reminders

Figure 2 shows the average standardized test scores on the full test across the three levels of the second factor that we manipulated. An analysis of variance with Bonferroni correction indicates that subjects' performance on the test significantly varied between groups (F $=15.11$, $p < .0005$). The result supports hypothesis H2a. On average, participants who received the visual reminder outperformed the control group by .18 standard deviations ($p < .0005$) and their peers exposed to the textual reminder by .11 standard deviations ($p = .005$). The latter group outperformed the control group by .08 standard deviations, although the p-value associated with this effect is .058, thus slightly above the standard significance threshold of .05. The findings support hypothesis H2b.

The interaction of defaults with reminders

We now test whether our two experimental factors interacted with each other. The left most panel in Figure 3 displays the difference in performance between subjects who took the regular test and their counterparts in the default (pre-populated) condition, separately for each of the three levels of the reminder manipulation. Relative to the control group (solid line), the decline in performance caused by the default answers is .20 standard deviation greater ($p = .004$) for subjects who were exposed to the textual reminder right before taking the test. The steeper inclination of the dashed line in the left most panel of Figure 6 indicates that the textual reminder exacerbated the negative impact of the default treatment on the overall test score. The decline in performance was not different between subjects in the control and the visual reminder arms at the usual statistical levels ($p = .145$).

Limited to the incorrect answers that might be pre-populated, the incorrect pre-populated panel of Figure 3 shows that the decline in performance was .17 standard deviations larger ($p = .010$) for those who received the textual reminder relative to the control group. As apparent from the parallel lines in the same panel, the default effect did not differ between the control and the visual reminder groups ($p = .449$).
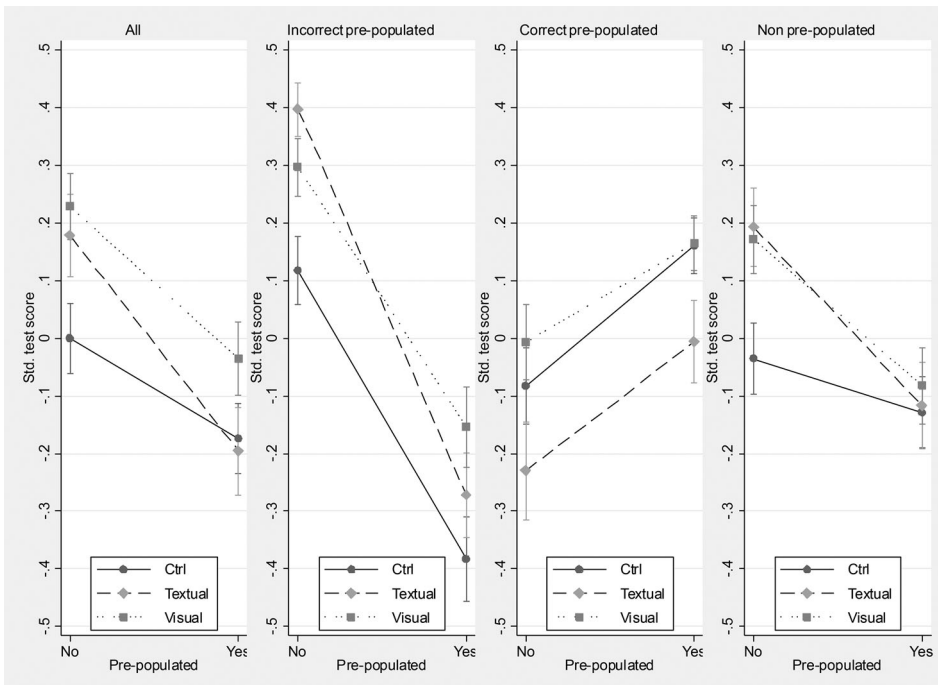
Figure 3. Interaction between defaults and reminders on the standardized score on the test, by typology of answer.
Note: Estimates from the linear regression model that underlies the figure are reported in Appendix B, Table B4.

The lack of divergence between the lines in the correct pre-populated panel of Figure 3 shows that the increase in performance caused by the inclusion of correct default answers did not significantly vary across types of reminders.

Instead, the interaction between our two factors was significant with regard to the standardized score on the answers that could not be pre-populated. As shown right most panel of Figure 3, relative to the control group, the loss in performance due to the default effect was .22 standard deviation greater ($p = .002$) under the textual reminder and .16 standard deviations larger for professionals exposed to the visual reminder ($p = .012$).

Taken as a whole, our interaction analyses suggest that reminders tend to be more effective in the absence of default rather than when defaults are present. In particular, the usefulness of textual reminders seems to be muted by defaults. Quite unexpectedly, healthcare workers who took our test found no benefit from reminders in the section of the test that was not directly affected by defaults, neither correct nor incorrect.

Discussion and implications

We contribute to nascent behaviorally focused research in the public sector (Battaglio et al., 2019; Cantarelli et al., 2020; Grimmelikhuijsen et al., 2017) by investigating two micro-level mechanisms, namely, defaults and reminders. Public organizations and their managers can leverage both mechanisms to boost their employees' knowledge of guidelines, ultimately encouraging compliance and helping to solve the implementation gap puzzle. In particular, our study provides testable hypotheses and a replicable research design, thus helping behavioral human resources flourish as a middle range theory in public management.

"Defaults are ubiquitous and powerful" (Thaler & Sunstein 2008, 92) and come in many forms in the context of public bureaucracies. Adherence to "the way things have always been done"

within a team or organization, for instance, may constitute a powerful default. Another example is the accumulation of errors within outdated procedures and regulations over time. Along the same lines, the reuse of an old form that includes outdated information may act as an incorrect default for decisions in the present. In general, practices that become consolidated habits, even if not supported by cutting-edge evidence, act as defaults that direct employees' conduct. These examples highlight the use of defaults as levers for modifying behaviors and boosting relevance. For example, Chiu and colleagues (2018) argue that pre-populating digital prescriptions with an untenably high number of pills, thus setting an incorrect default, may lead to an excessive pre-scribing of opioids, ultimately exacerbating the opioid epidemic. Conversely, lowering the pre-populated number of pills—setting a correct default—decreases the amount of opioids prescribed. Because our results indicate that the harm from incorrect defaults outweigh the benefits of correct defaults, public organizations and their managers might devote greater effort to detecting and removing incorrect defaults. Our results indicate that if choice architects require a specific answer from individuals, they should consider setting said answer as the default. If, on the other hand, their aim is to encourage individuals to reflect and work out answers on their own they should avoid setting any defaults. These are actionable findings in the form of default nudges and can easily be amended and disclosed without precluding psychological reticence from those affected (Sunstein et al., 2019).

Like defaults, reminders are also interventions that tend to pervade public bureaucracies (Andersen & Hvidman, 2021). The means by which information is presented is an increasingly salient topic for decisions and performance in public bureaucracies, especially healthcare organizations at the front lines of unforeseen events—such as the COVID-19 pandemic—who rely on cues to inform their staff. In particular, we test the impact of information—textual or visual—which should presumably be irrelevant because its knowledge is required by law. Our findings convey the superiority of visual aids—relative to textual reminders—as a means for public managers to possibly bridge the guidelines implementation gap in their organizations. This is an especially relevant finding given the lack of research on the use of nudges in the form of reminders in hand hygiene research (Nagtegaal et al., 2019)—an imperative area of research for the ongoing global pandemic. Employing visual reminders, particularly in the case of healthcare professionals, is critical for easing information diffusion at point-of-care (Nagtegaal et al., 2019). Such visual cues have the potential to improve information diffusion and compliance with policy guidelines (John & Blume, 2018; Michie & Lester, 2005).

Our interaction analyses strengthen the findings, especially the impact of reminders and reminder formats, which are moderated by the opportunity of setting defaults. Practitioners in public institutions might be circumspect in their reliance on reminders and consider the useful-ness of defaults when both tools are available as mechanisms for desired behaviors. Our novel randomized design allows for a process of early participation among affected groups (healthcare professionals), public accountability, careful deliberation, and transparency (Sunstein et al., 2019). By employing a 'test-learn-adapt-share' approach, we emulate successful practices increasingly promoted by leading behavioral policy labs (Haynes et al. 2012) in a public management setting. Our analyses of performance on the test items that were not pre-populated underscore the separate and joint impact of defaults and reminders as tools to enhance compliance among public employees. In fact, reminders were of no help in counterbalancing the effect of defaults, both correct and incorrect, on the score of answers that could not be pre-populated.

Research on defaults and reminders that focuses on public employee appreciation of guidelines may further our understanding of the implementation gap in public policy and administration. Policy design, on the one hand, and policy evaluation, on the other hand, are two phases of the policy cycle that tend to capture the attention of most politicians, the media, and the public. Unfortunately, policy implementation is often neglected in general public administration literature and more specifically public management (O'Toole, 2017), where "practical challenges of effective

policy implementation loom, if anything, larger than ever" (376). Indeed, well-designed policies often fail to deliver on their promises and fall short of expectations due to execution shortcomings. The use of guidelines—widely adopted in governments globally in a broad array of policy areas—does not seem to be extraneous to such execution shortcomings. This phenomenon, known as the implementation gap, has long drawn attention from public administration and management scholars alike (Ongaro & Valotti, 2008).

Our work on guidelines implementation in healthcare settings connects very naturally and nicely with latest developments in the study of judgment errors (Kahneman, Sibony, & Sunstein, 2021). Specifically, Kahneman et al. (2021) suggest that guidelines in medicine are a viable tool to simultaneously reduce the average of errors (i.e., systematic biases) and the variability of errors (i.e., noise). Guidelines for the provision of health services can take the form of either rules—where any room for decision makers' discretion is minimal—or standards—where decision makers' discretion is larger though not full. Taken together, this evidence reinforces the relevance of our study to further our understanding about how the architecture of choice can encourage the implementation of guidelines in the health domain.

Overall, the findings provide several takeaways in response to recent calls to apply behavioral insights toward the goal of making public policies work better across domains (Organization for Economic Cooperation and Development, 2017). Our work bolsters the growing body of behaviorally focused research interested in fine-tuning the knowledge and implementation of guidelines. Most notably, we employ a large sample of public employees to explore behavioral issues among public employees—a key component of measurement theory (see Meier & O'Toole, 2012). Additionally, the randomized nature of our research design offers practitioners causal evidence that is applicable to early stages of the policy cycle. This is especially useful to healthcare professionals at the front-line of the COVID-19 pandemic who are continuously modifying behaviors to cope with the ongoing crisis.

Limitations

We certainly acknowledge that our study is not without limitations. Therefore, we urge public managers and policy makers to interpret and use our findings considering its inherent shortcomings. In particular, our randomized controlled trial shares common weaknesses with studies of the same kind. In fact, as for most experimental studies published in top tier journals, our experiment scores high on internal validity at the cost of lower external validity (Belle and Cantarelli 2018). However, this concern is mitigated by the nature of our sample, which consists of real healthcare workers participating in a routine survey administered on behalf of their employers (Mele et al., 2021). Furthermore, recent scholarship suggests that individuals "by and large, […] address a plausible, hypothetical problem in much the same way that [they] tackle a real one." (Kahneman, Sibony, and Sunstein 2021:48).

First, concerns about the external validity of our results are understandable. Large-scale survey experiments with real public employees engaged in tasks strictly related to their daily job activities score higher on external validity relative, for instance, to laboratory experiments. Nonetheless, whether and to what extent our findings would be replicated in more naturally occurring settings is yet to be demonstrated. For example, the generalizability of our results that are based on the scores of an online test to guidelines compliance in actual behaviors cannot be taken for granted. Along the same lines, whether and how our findings derived from a potentially low-stake task would hold in the context of higher-stake tasks is still an open question. To address this limitation, natural field experiments are superior to survey experiments in allowing better generalizability of results. Furthermore, the degree to which the pattern of our results is generalizable across public professions, guidelines, settings, treatments, and operations is yet to be tested. For instance, novel work may better equalize the content included in textual and visual reminders. It is

important to also note that the leaflet we designed with the partner organization's healthcare staff includes text alongside images. As a result, our experimental design only allows estimating the causal impact of visual reminders that contain both images and text, not the effect of purely pictorial reminders. While in our field experiment the use of purely pictorial memos would not have been feasible and would have compromised realism, future studies could verify the effect of text-less reminders in different research contexts in which this type of aid is usable. In general, we are convinced that there is no perfect operationalization and that ours shares the intrinsic limits of any alternative operationalization. Every operationalization, including the one we have used, implies tradeoffs between precision and realism, that is, between construct and external validity. For instance, in designing our visual reminder treatment, we prioritized realism over accuracy. The motivation behind this choice was to generate results that could be more easily implemented in the field. Along the same lines, future studies may be designed in which correct and incorrect answers are not mutually exclusive. In our study, incorrectly and correctly pre-populated responses are systematically linked to different items. This could not be otherwise because the same circumstance cannot be both true and false. On the one hand, readers are certainly warned that the direct comparison between correct and incorrect defaults must be interpreted in light of the fact that correctly and incorrectly pre-populated answers necessarily relate to different items. On the other hand, readers can be reassured that this inherent feature does not invalidate the comparison of each default type—correct or incorrect—and the control group. Testing the interaction between defaults and ease of identification of a situation as requiring the use of gloves was beyond the scope of our study. Nevertheless, studying the interaction of defaults with difficulty in mapping circumstances seems a natural extension of our RCT that is worth pursuing.

Although our factorial design is best suited to estimate the average treatment effects of our two factors, as well as their interaction, it does not explain the causal mechanism through which the effects we observe come about. In other words, whereas our randomized controlled trial allows making causal claims about the impact of defaults and reminders on workers' knowledge of guidelines, it does not illuminate the underlying chain reaction leading from our manipulations to their outcomes. This paves the way for future research that uses more sophisticated experimental designs, such as the parallel encouragement design, and triangulates quantitative and qualitative evidence to gain a richer understanding of the *how* in addition to the *what* (Mele & Belardinelli, 2019). Such work might, for instance, shed light on whether and to what degree the impact of defaults unfolds by reducing individuals' reflection about the correct answer.

Conclusion

Micro-level aspects of cognitive and decision processes are increasingly a germane topic for public policy and management scholars. These processes are subject to behavioral dissonance during the policy process, especially the implementation stage where policy actors are often "trapped in unproductive patterns of interaction" (O'Toole, 2004, p. 309). Nascent scholarship has appealed for the use of insights from psychology to address pressing public management issues by leveraging micro-level mechanisms to promote desired attitudes and behaviors among civil servants (Battaglio et al., 2019; Cantarelli et al., 2020). Responding to those recent calls, we have defined and operationalized our key study constructs based on well-established behavioral science research on defaults and reminders. The rationale behind this choice is twofold. Drawing on established knowledge not only maximizes the validity of our inference but, at the same time, it also strengthens the scientific dialogue between ours and allied disciplines.

The research presented here confirms the importance of employing the right types of nudges. Regarding defaults, this means prioritizing the removal of incorrectly pre-populated options. With respect to reminders, we demonstrate the advantage of using visual cues, as opposed to textual, in the framing of choice architecture. The findings from our interaction analyses bolster

these themes. Indeed, setting defaults overshadowed the impact of reminders and reminder for-mats. In conclusion, our research may represent a solid way of tackling public management issues that "adopt a more solution-oriented approach, starting first with a practical problem and then asking what theories (and methods) must be brought to bear to solve it" (Watts, 2017:1).

Notes on contributors

Nicola Belle (nicola.belle@santannapisa.it) is an associate professor at Scuola Superiore Sant'Anna (Management and Healthcare Laboratory, Institute of Management). His research focuses on experimental and behavioral public administration.

Paola Cantarelli (paola.cantarelli@santannapisa.it) is an assistant professor at Scuola Superiore Sant'Anna (Management and Healthcare Laboratory, Institute of Management). Her research focuses on behavioral human resource management and work motivation in mission-driven organizations.

Chiara Barchielli is a PhD in Management Innovation, Sustainability and Healthcare at Scuola Superiore Sant'anna and a research nurse for the Azienda USL Toscana Centro. Her research focuses on healthcare delivery models and healthcare management.

Paul Battaglio, PhD., is the former editor of Public Administration Review and the Review of Public Personnel Administration. His research interests include behavioral public administration, human resource management, organization theory and behavior, and comparative public policy.

References

Allcott, H., and T. Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104(10):3003–37. doi: 10.1257/aer.104.10.3003.

Altmann, S., and C. Traxler. 2014. "Nudges at the Dentist." *European Economic Review* 72:19–38. doi: 10.1016/j.euroecorev.2014.07.007.

Andersen, S. C., and U. Hvidman. 2021. "Can Reminders and Incentives Improve Implementation within Government? Evidence from a Field Experiment." *Journal of Public Administration Research and Theory* 31(1): 234–49. doi: 10.1093/jopart/muaa022.

Baekgaard, M., N. Belle, S. Serritzlew, M. Sicilia, and I. Steccolini. 2019. "Performance information in Politics: How Framing, Format, and Rhetoric Matter to Politicians' Preferences." *Journal of Behavioral Public Administration* 2(2):1–12. doi: 10.30636/jbpa.22.67.

Battaglio, R. P., P. Belardinelli, N. Bellé, and P. Cantarelli. 2019. "Behavioral Public Administration Ad Fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations." *Public Administration Review* 79(3):304–20. doi: 10.1111/puar.12994.

Bauer, M. S., L. Damschroder, H. Hagedorn, J. Smith, and A. M. Kilbourne. 2015. "An Introduction to Implementation Science for the Non-Specialist." *BMC Psychology* 3(1):32. doi: 10.1186/s40359-015-0089-9.

Belle, N., and P. Cantarelli. 2018. "Randomized experiments and Reality of Public and Nonprofit Organizations: Understanding and Bridging the Gap." *Review of Public Personnel Administration* 38(4):494–511.

Belle, N., P. Cantarelli, and P. Belardinelli. 2018. "Prospect Theory Goes Public: Experimental Evidence on Cognitive Biases in Public Policy and Management Decisions." *Public Administration Review* 78(6):828–40. doi: 10.1111/puar.12960.

Bhanot, S. P., and E. Linos. 2020. "Behavioral Public Administration: Past, Present, and Future." *Public Administration Review* 80(1):168–71. doi: 10.1111/puar.13129.

Bhargava, S., and D. Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105(11):3489–529. doi: 10.1257/aer.20121493.

Blumenthal-Barby, J. S., and H. Krieger. 2015. "Cognitive Biases and Heuristics in Medical Decision Making: A Critical Review Using a Systematic Search Strategy." *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 35(4):539–57. doi: 10.1177/0272989X14547740.

Bresciani, S., and M. J. Eppler. 2009. "The Benefits of Synchronous Collaborative Information Visualization: Evidence from an Experimental Evaluation." *IEEE Transactions on Visualization and Computer Graphics* 15(6): 1073–80. doi: 10.1109/TVCG.2009.188.

Brown, Louise, Julianne Munro, and Suzy Rogers. 2019. "Use of Personal Protective Equipment in Nursing Practice." *Nursing standard (Royal College of Nursing (Great Britain): 1987)* 34(5):59–66. doi: 10.7748/ns.2019.e11260.

Cantarelli, P., N. Belle, and P. Belardinelli. 2020. "Behavioral Public HR: Experimental Evidence on Cognitive Biases and Debiasing Interventions." *Review of Public Personnel Administration* 40(1):56–81. doi: 10.1177/0734371X18778090.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick. 2009. "Optimal Defaults and Active Decisions." *The quarterly Journal of Economics* 124(4):1639–74. doi: 10.1162/qjec.2009.124.4.1639.

Center for Disease Control and Prevention. 2019. Guideline for Isolation Precautions: Preventing Transmission of Infectious Agents in Healthcare Settings. https://www.cdc.gov/infectioncontrol/pdf/guidelines/isolation-guidelines-H.pdf [Last accessed on August 28, 2020].

Chiu, A., S. Jean, R. A. Hoag, J. R. Freedman-Weiss, M. Healy, J. M, and Pei, K. Y. 2018. "Association of Lowering Default Pill Counts in Electronic Medical Record Systems with Postoperative Opioid Prescribing." *JAMA surgery* 153(11):1012–9. doi: 10.1001/jamasurg.2018.2083.

Choueiki, A. 2016. "Behavioral Insights for Better Implementation in Government." *Public Administration Review* 76(4):540–1. doi: 10.1111/puar.12594.

Dale, A., and A. Strauss. 2009. "Don't Forget to Vote: Text Message Reminders as a Mobilization Tool." *American Journal of Political Science* 53(4):787–804. doi: 10.1111/j.1540-5907.2009.00401.x.

DellaVigna, S., and E. Linos. 2022. "RCTs to Scale: Comprehensive Evidence from Two Nudge Units." *Econometrica*, 90(1):81–116. doi: 10.3982/ECTA18709.

Exworthy, M., and M. Powell. 2004. "Big Windows and Little Windows: Implementation in the 'Congested State'." *Public Administration* 82(2):263–81.

Fonseca, M. A., and S. B. Grimshaw. 2017. "Do Behavioral Nudges in Prepopulated Tax Forms Affect Compliance? Experimental Evidence with Real Taxpayers." *Journal of Public Policy & Marketing* 36(2):213–26.

Fryer, K., J. Antony, and S. Ogden. 2009. "Performance Management in the Public Sector." *International Journal of Public Sector Management* 22(6):478–98. doi: 10.1108/09513550910982850.

Green, D. P., and A. S. Gerber. 2019. *Get Out the Vote: How to Increase Voter Turnout*. Washington, D.C.: Brookings Institution Press.

Grimmelikhuijsen, S., S. Jilke, A. L. Olsen, and L. Tummers. 2017. "Behavioral Public Administration: Combining Insights from Public Administration and Psychology." *Public Administration Review* 77(1):45–56. doi: 10.1111/puar.12609.

Haynes, L. C., D. P. Green, R. Gallagher, P. John, and D. J. Torgerson. 2013. "Collection of Delinquent Fines: An Adaptive Randomized Trial to Assess the Effectiveness of Alternative Text Messages." *Journal of Policy Analysis and Management* 32(4):718–30. doi: 10.1002/pam.21717.

Haynes, L., O. Service, B. Goldacre, and D. Torgerson. 2012. *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials | Cabinet Office*. Technical Report. Cabinet Office Behavioural Insights Team, UK. https://researchonline.lshtm.ac.uk/id/eprint/201256

Hildon, Z., D. Allwood, and N. Black. 2012. "Making Data More Meaningful: Patients' Views of the Format and Content of Quality Indicators Comparing Health Care Providers." *Patient education and Counseling* 88(2):298–304. doi: 10.1016/j.pec.2012.02.006.

Imai, K., D. Tingley, and T. Yamamoto. 2013. "Experimental Designs for Identifying Causal Mechanisms: Experimental Designs for Identifying Causal Mechanisms." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176(1):5–51. doi: 10.1111/j.1467-985X.2012.01032.x.

Isett, K. R., and D. M. Hicks. 2018. "Providing Public Servants What They Need: Revealing the "Unseen" through Data Visualization." *Public Administration Review* 78(3):479–85. doi: 10.1111/puar.12904.

Jilke, S. 2015. "Choice and Equality: Are Vulnerable Citizens Worse off after Liberalization Reforms?" *Public Administration* 93(1):68–85. doi: 10.1111/padm.12102.

Jilke, S., G. G. Van Ryzin, and S. Van de Walle. 2016. "Responses to Decline in Marketized Public Services: An Experimental Evaluation of Choice Overload." *Journal of Public Administration Research and Theory* 26(3):421–32. doi: 10.1093/jopart/muv021.

John, P., and T. Blume. 2018. "How Best to Nudge Taxpayers? The Impact of Message Simplification and Descriptive Social Norms on Payment Rates in a Central London Local Authority." *Journal of Behavioral Public Administration* 1(1):1–11.

Johnson, E. J., and D. Goldstein. 2003. "Do Defaults Save Lives?" *Science* 302(5649):1338–9.

Kahneman, D., O. Sibony, and C. R. Sunstein. 2021. *Noise: A Flaw in Human Judgment*. New York: Little, Brown Spark.

Kang, JaHyun, John M. O'Donnell, Bonnie Colaianne, Nicholas Bircher, Dianxu Ren, and Kenneth J. Smith. 2017. "Use of Personal Protective Equipment among Health Care Personnel: Results of Clinical Observations and Simulations." *American Journal of Infection Control* 45(1):17–23. doi: 10.1016/j.ajic.2016.08.011.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62(12):3393–411. doi: 10.1287/mnsc.2015.2296.

McDaniels, T. L. 1996. "The Structured Value Referendum: Eliciting Preferences for Environmental Policy Alternatives." *Journal of Policy Analysis and Management* 15(2):227–51. doi: 10.1002/(SICI)1520-6688(199621)15:2<227::AID-PAM4>3.0.CO;2-L.

Meier, K. J., and L. J. O'Toole. 2012. "Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships." *Journal of Public Administration Research and Theory* 23(2):429–56.

Mele, V., and P. Belardinelli. 2019. "Mixed Methods in Public Administration Research: Selecting, Sequencing, and Connecting." *Journal of Public Administration Research and Theory* 29(2):334–47. doi: 10.1093/jopart/muy046.

Mele, V., N. Bellé, and M. Cucciniello. 2021. "Thanks, but No Thanks. Preferences towards Teleworking Colleagues in Public Organizations." *Journal of Public Administration Research and Theory* 31(4):790–805.

Michie, S., and K. Lester. 2005. "Words Matter: Increasing the Implementation of Clinical Guidelines." *Quality & Safety in Health Care* 14(5):367–70. doi: 10.1136/qshc.2005.014100.

Moseley, A., and G. Stoker. 2015. "Putting Public Policy Defaults to the Test: The Case of Organ Donor Registration." *International Public Management Journal* 18(2):246–64. doi: 10.1080/10967494.2015.1012574.

Moynihan, D. P., and S. Lavertu. 2012. "Cognitive Biases in Governing: Technology Preferences in Election Administration." *Public Administration Review* 72(1):68–77. doi: 10.1111/j.1540-6210.2011.02478.x.

Münscher, R., M. Vetter, and T. Scheuerle. 2016. "A Review and Taxonomy of Choice Architecture Techniques." *Journal of Behavioral Decision Making* 29(5):511–24. doi: 10.1002/bdm.1897.

Nagtegaal, R., L. Tummers, M. Noordegraaf, and V. Bekkers. 2019. "Nudging Healthcare Professionals towards Evidence-Based Medicine: A Systematic Scoping Review." *Journal of Behavioral Public Administration* 2(2):1–20. doi: 10.30636/jbpa.22.71.

Nevo, I., M. Fitzpatrick, R.-E. Thomas, P. A. Gluck, J. D. Lenchus, K. L. Arheart, and D. J. Birnbach. 2010. "The Efficacy of Visual Cues to Improve Hand Hygiene Compliance." *Simulation in Healthcare* 5(6):325–31.

Olsen, A. L. 2017. "Human Interest or Hard Numbers? Experiments on Citizens' Selection, Exposure, and Recall of Performance Information." *Public Administration Review* 77(3):408–20. doi: 10.1111/puar.12638.

Ongaro, E., and G. Valotti. 2008. "Public Management Reform in Italy: Explaining the Implementation Gap." *International Journal of Public Sector Management* 21(2):174–204. doi: 10.1108/09513550810855654.

Organization for Economic Cooperation and Development. 2017. "Behavioural Insights and Public Policy Lessons from around the World."

O'Toole, L. J. Jr. 2004. "The Theory–Practice Issue in Policy Implementation Research." *Public Administration* 82(2):309–29.

O'Toole, L. J. 2017. "Implementation for the Real World." *Journal of Public Administration Research and Theory* 27(2):376–9.

Patel, M. S., S. C. Day, S. D. Halpern, C. W. Hanson, J. R. Martinez, S. Honeywell, and K. G. Volpp. 2016. "Generic Medication Prescription Rates after Health System-Wide Redesign of Default Options within the Electronic Health Record." *JAMA internal Medicine* 176(6):847–8. doi: 10.1001/jamainternmed.2016.1691.

Rahman, A. A., Y. B. Adamu, and P. Harun. 2017. "Review on Dashboard Application from Managerial Perspective." 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), 1–5. doi: 10.1109/ICRIIS.2017.8002461.

Riccucci, N. M., M. K. Meyers, I. Lurie, and J. S. Han. 2004. "The Implementation of Welfare Reform Policy: The Role of Public Managers in Front-Line Practices." *Public Administration Review* 64(4):438–48. doi: 10.1111/j.1540-6210.2004.00390.x.

Sanders, M., V. Snijders, and M. Hallsworth. 2018. "Behavioural Science and Policy: Where Are we Now and Where Are we Going?" *Behavioural Public Policy* 2(2):144–67.

Saposnik, G., D. Redelmeier, C. C. Ruff, and P. N. Tobler. 2016. "Cognitive Biases Associated with Medical Decisions: A Systematic Review." *BMC medical Informatics and Decision Making* 16(1):138. doi: 10.1186/s12911-016-0377-1.

Shpaizman, I. 2017. "Policy Drift and Its Reversal: The Case of Prescription Drug Coverage in the United States." *Public Administration* 95(3):698–712. doi: 10.1111/padm.12315.

Siciliano, M. D., N. M. Moolenaar, A. J. Daly, and Y.-H. Liou. 2017. "A Cognitive Perspective on Policy Implementation: Reform Beliefs, Sensemaking, and Social Networks." *Public Administration Review* 77(6):889–901. doi: 10.1111/puar.12797.

Slingsby, A., J. Dykes, J. Wood, and R. Radburn. 2014. "Designing an Exploratory Visual Interface to the Results of Citizen Surveys." *International Journal of Geographical Information Science* 28(10):2090–125. doi: 10.1080/13658816.2014.920845.

Sunstein, C. R., L. A. Reisch, and M. Kaiser. 2019. "Trusting Nudges? Lessons from an International Survey." *Journal of European Public Policy* 26(10):1417–43.

Thaler, R. 2017. Nobel Prize Lecture "From Cashews to Nudges: The Evolution of Behavioral Economics". https://www.youtube.com/watch?v=ej6cygeB2X0 [Last accessed on August 28, 2020].

Thaler, R. H. 2015. "Misbehaving: The Making of Behavioral Economics."

Thaler, R. H., and S. Benartzi, S. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of political Economy* 112(S1):S164–87.

Thaler, R. H., and C. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

Thomann, E. 2018. "Donate Your Organs, Donate Life! Explicitness in Policy Instruments." *Policy Sciences* 51(4): 433–56.

van Dijk, W. W., S. Goslinga, B. W. Terwel, and E. van Dijk. 2020. "How Choice Architecture Can Promote and Undermine Tax Compliance: Testing the Effects of Prepopulated Tax Returns and Accuracy Confirmation." *Journal of Behavioral and Experimental Economics* 87:101574.

Verbeek, Jos H., Blair Rajamaki, Sharea Ijaz, Riitta Sauni, Elaine Toomey, Bronagh Blackwood, Christina Tikka, Jani H. Ruotsalainen, and F. Selcen Kilinc Balci. 2020. "Personal Protective Equipment for Preventing Highly Infectious Diseases Due to Exposure to Contaminated Body Fluids in Healthcare Staff." *The Cochrane Database of Systematic Reviews* 5(5):CD011621. doi: 10.1002/14651858.CD011621.pub3.

Watts, D. J. 2017. "Should Social Science Be More Solution-Oriented?" *Nature Human Behaviour* 1(1):1–5. doi: 10.1038/s41562-016-0015.

World Health Organization. 2009. A Guide to the Implementation of the WHO Multimodal Hand Hygiene Improvement Strategy https://www.who.int/gpsc/5may/Guide_to_Implementation.pdf?ua=1 [Last accessed on August 28, 2020].

World Health Organization. 2016. Personal Protective Equipment for Use in a Filovirus Disease Outbreak. Rapid advice guideline https://apps.who.int/iris/bitstream/handle/10665/251426/9789241549721-eng.pdf [Last accessed on August 28, 2020].

Appendix A

In the list below, please select the situations in which the use of gloves is correct according to the indications of the Executive Decree 19154/2017.
- ☐ Procedures in sterile field
- ☐ Contact with inanimate objects
- ☐ Predictable contact with blood and / or biological fluids, including injured skin and mucous membranes, excluding sweating
- ☐ Predictable contact with hazardous chemicals, including drugs for which a hazard to the operator is highlighted - e.g. antiblastic -, sharp materials, ionizing radiation
- ☐ In the event of the spread of spores
- ☐ Contact with antibiotics
- ☐ Assistance to patients known or suspected of being infected or colonized by epidemiologically important pathogenic microorganisms transmissible by contact

Figure A1. Scenario in the *no default* (i.e. *no pre-populated*) treatment.

In the list below, please select the situations in which the use of gloves is correct according to the indications of the Executive Decree 19154/2017. Deselect any situations that are incorrectly pre-selected.
- ☐ Procedures in sterile field
- ☒ Contact with inanimate objects
- ☒ Predictable contact with blood and / or biological fluids, including injured skin and mucous membranes, excluding sweating
- ☐ Predictable contact with hazardous chemicals, including drugs for which a hazard to the operator is highlighted - e.g. antiblastic -, sharp materials, ionizing radiation
- ☐ In the event of the spread of spores
- ☒ Contact with antibiotics
- ☒ Assistance to patients known or suspected of being infected or colonized by epidemiologically important pathogenic microorganisms transmissible by contact

Figure A2. Scenario in the *default* (i.e. *pre-populated*) treatment.

In the list below, please select the situations in which the use of gloves is correct according to the indications of the Executive Decree 19154/2017.

Figure A3. Scenario in the *no reminder* treatment.

Please read carefully the excerpt of the Executive Decree 19154/2017 that is reported below. It indicates the situations in which the use of gloves is correct.

> **2. Object and operating methods**
> **2.1 Purpose**
> These indications for use originate from the following general considerations, according to which the objectives of interest consist in achieving the best level of safety and biocompatibility, both for the patient and for the healthcare worker.
> 1. Use gloves <u>only</u> in the following situations:
>   a. Procedures in sterile field (patient's safety)
>   b. Predictable contact with blood and / or biological fluids, including injured skin and mucous membranes, excluding sweating (healthcare worker's safety)
>   c. Assistance to patients known or suspected of being infected or colonized by epidemiologically important pathogenic microorganisms transmissible by contact (healthcare worker and other patients' safety)
>   d. Predictable contact with hazardous chemicals, including drugs for which a hazard to the operator is highlighted - e.g. antiblastic -, sharp materials, ionizing radiation (healthcare worker's safety)
> 2. Reserve the use of latex gloves exclusively for the surgical field, with the exception of paths, environments and latex free operating rooms, and resorting to gloves of different materials in all other uses.
> 3. Eliminate the use of powder-lubricated gloves.

Figure A4. Scenario in the *textual reminder* treatment.

Please look carefully at the leaflet below. It illustrates the situations in which the use of gloves is incorrect and the situations in which it is correct according to the Executive Decree 19154/2017.

Handling of documents	Procedures in sterile field (patient's safety)
Handling of test tubes, bags, containers, that do not contain biological fluids and/or chemicals	Predictable contact with blood and / or biological fluids, including injured skin and mucous membranes, excluding sweating (healthcare worker's safety)
Therapy for oral administration, except for drugs for which a hazard to the operator is highlighted - e.g. antiblastic	
Medical examination and assistance to the patient in which contact with blood, fluids, and potentially infectious biological agents is not foreseeable	Assistance to patients known or suspected of being infected or colonized by epidemiologically important pathogenic microorganisms transmissible by contact (healthcare worker and other patients' safety)
Making the empty bed	Predictable contact with hazardous chemicals, including drugs for which a hazard to the operator is highlighted - e.g. antiblastic -, sharp materials, ionizing radiation (healthcare worker's safety)
Washing of hands and face of bedridden patients	
Distribution of food trays	
Changing drips, except for drugs for which a hazard to the operator is highlighted - e.g. antiblastic	
Transport of patients with wheelchair or bed	
Measurement of body temperature, blood pressure, and heart rate	
Aerosol therapy and oxygen therapy	

Figure A5. Scenario in the *visual reminder* treatment.

Appendix B

Table B1. Average unstandardized scores with standard deviations in parenthesis, by type of answers, by experimental arm.

		Type of answers			
		Full test (possible range was from -3 to +4)	Incorrect answers that might be pre-populated (possible range was from 0 to +2)	Correct answers that might be pre-populated (possible range was from 0 to +2)	Answers that were not pre-populated (possible range was from -1 to +2)
Experimental arm	No default (no pre-populated) n = 2,649	2.71 (1.19)	.34 (.62)	1.81 (.51)	1.24 (.65)
	Default (pre-populated) n = 2,654	2.24 (1.41)	.76 (.86)	1.91 (.40)	1.09 (.70)
	No reminder n = 1,768	2.33 (1.24)	.65 (.81)	1.88 (.41)	1.11 (.64)
	Textual reminder n = 1,768	2.49 (1.46)	.50 (.78)	1.81 (.55)	1.19 (.75)
	Visual reminder n = 1,767	2.58 (1.24)	.49 (.74)	1.90 (.40)	1.19 (.65)

Table B2. Results of a linear regression model estimating the average standardized score as a function of default treatments, by type of answers.

	Full test				Incorrect answers that might be pre-populated				Correct answers that might be pre-populated				Answers that were not pre-populated			
	β	RSE	t	p	β	RSE	t	p	β	RSE	t	p	β	RSE	t	p
No default (ref.)																
Default	-.27	.03	−9.94	.000	-.54	.03	−20.44	.000	.21	.03	7.79	.000	-.22	.03	−8.02	.000
Constant	.14	.02	7.25	.000	.27	.02	17.48	.000	-.11	.02	−4.93	.000	.11	.02	5.89	.000
Obs.	5,303				5,303				5,303				5,303			
F	98.75				417.85				60.67				64.37			
p	.000				.000				.000				.000			
R^2	.02				.07				.01				.01			
Root MSE	.99				.97				.99				.99			

Table B3. Results of a linear regression model estimating the average standardized score on the full test as a function of reminder treatments.

	β	RSE	t	p
No reminder (ref.)				
Textual reminder	.08	.03	2.26	.024
Visual reminder	.18	.03	5.92	.000
Constant	-.09	.02	−3.98	.000
Obs.	5,303			
F	17.60			
p	.000			
R^2	.01			
Root MSE	1.00			

Table B4. Results of a linear regression model estimating the interaction between defaults and reminders on standardized score on test, by type of answers.

	Full test				Incorrect answers that might be pre-populated				Correct answers that might be pre-populated				Answers that were not pre-populated			
	β	RSE	t	p	β	RSE	t	p	β	RSE	t	p	β	RSE	t	p
No default (ref.)																
Default	-.17	.04	-3.97	.000	-.50	.05	-10.41	.000	.24	.04	5.78	.000	-.09	.04	-2.10	.036
No reminder (ref.)																
Textual reminder	.18	.05	3.76	.000	.28	.04	7.31	.000	-.15	.06	-2.66	.008	.23	.05	4.87	.000
Visual reminder	.23	.04	5.39	.000	.18	.04	4.57	.000	.08	.05	1.59	.112	.21	.04	4.76	.000
Default * No reminder (ref.)																
Default * Textual reminder	-.20	.07	-2.91	.004	-.17	.07	-2.57	.010	-.02	.07	-.27	.785	-.22	.07	-3.17	.002
Default * Visual reminder	-.09	.06	-1.46	.145	.05	.07	.76	.449	-.07	.06	-1.21	.226	-.16	.06	-2.53	.012
Constant	-.00	.03	-.01	.989	.12	.03	3.91	.000	-.08	.03	-2.43	.015	-.04	.03	-1.12	.263
Obs.	5,303				5,303				5,303				5,303			
F	28.67				105.89				21.01				19.46			
p	.000				.000				.000				.000			
R^2	.02				.08				.02				.02			
Root MSE	.99				.96				.99				.99			