Keefer, Philip; Roseth, Benjamin; Santamaria, Julieth

Working Paper

General skills training for public employees: Experimental evidence on cybersecurity training in Argentina

IDB Working Paper Series, No. IDB-WP-1643

Provided in Cooperation with:
Inter-American Development Bank (IDB), Washington, DC

Suggested Citation: Keefer, Philip; Roseth, Benjamin; Santamaria, Julieth (2024) : General skills training for public employees: Experimental evidence on cybersecurity training in Argentina, IDB Working Paper Series, No. IDB-WP-1643, Inter-American Development Bank (IDB), Washington, DC, https://doi.org/10.18235/0013202

This Version is available at:
https://hdl.handle.net/10419/309120

WORKING PAPER Nº IDB-WP-1643

General Skills Training for Public Employees

Experimental Evidence on Cybersecurity Training in Argentina

Philip Keefer
Benjamin Roseth
Julieth Santamaria

Inter-American Development Bank
Institutions for Development Sector
Innovation in Citizen Services Division

October 2024

General Skills Training for Public Employees

Experimental Evidence on Cybersecurity Training in Argentina

Philip Keefer
Benjamin Roseth
Julieth Santamaria

Inter-American Development Bank
Institutions for Development Sector
Innovation in Citizen Services Division

October 2024

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Keefer, Philip.
General skills training for public employees experimental: evidence on cybersecurity training in Argentina / Philip Keefer, Benjamin Roseth, Julieth Santamaria.
p. cm. — (IDB Working Paper Series; 1643)
Includes bibliographical references.
1. Cyberterrorism-Argentina. 2. Phishing-Argentina. I. Roseth, Benjamin. II. Santamaria, Julieth. III. Inter-American Development Bank. Innovation in Citizen Services Division. IV. Title. V. Series.
IDB-WP-1643

JEL Codes: N46
Keywords: cybersecurity, cyberterrorism, phishing

http://www.iadb.org

Abstract

Cyberattacks have risen to become one of the most critical global risks. Despite increasing investments to combat cyberattacks, there remains a significant, often unnoticed vulnerability: employees. Previous literature reveals that over two-thirds of cyberattacks within organizations result from employee negligence. While strengthening cybersecurity through employee training is essential, traditional methods often fall short. In this study, we tested different approaches to reduce risk exposure to phishing, one of the most common types of cyberattacks, focusing on a sector and context unaddressed by previous literature: the public sector in a developing country (Argentina). We randomly allocated 1,918 public servants to a control group and two treatment groups to compare the effectiveness of online training—commonly used to promote behavior changes on ancillary workplace topics such as ethics, discrimination, and data protection—versus a "learning-by-doing" approach, which involved sending repeated phishing emails followed by educational emails. Our findings indicate that the learning-by-doing approach is superior for enhancing phishing email detection, resulting in fewer phishing emails opened, fewer clicks on phishing links, and improved reporting of suspicious emails. This strategy is particularly effective among permanent public officials compared to contractors, as well as among female employees. These findings not only inform organizational cybersecurity practices but also have broader implications for influencing employee behavior on other important workplace topics.

1. Introduction

Public institutions and enterprises recognize the importance of equipping employees with general skills and promoting behaviors that enhance productivity and reduce costs. These topics are often peripheral to employees' primary tasks but are closely related to institutional norms, such as ethics, discrimination, and data protection. To address these areas, institutions often offer on-the-job training programs, usually in the form of mandatory courses. However, these training programs often face challenges related to employee engagement and effectiveness. Employee disinterest in acquiring these skills often results in minimal learning and limited behavior change. This disengagement is primarily driven by two factors: first, employees may not see these skills as beneficial for their resumes or future career prospects; second, they may fail to understand how critical these competencies are to the organization's success. To address this issue, institutions must implement effective strategies to motivate and engage employees, helping them develop these skills and adopt key behaviors.

As online tools become increasingly integral to job roles and cyberattacks grow more frequent and sophisticated, cybersecurity has become a critical area for employee training, despite "cybersecurity" being mentioned in relatively few job descriptions. The World Economic Forum (WEF) ranks cyberattacks among the top seven global risks, with significant financial impacts worldwide (WEF, 2022). In 2021, global cybercrime costs surpassed $6 trillion (Cybercrime Magazine, 2017). Recent attacks in countries like Costa Rica, Peru, and Argentina underscore the vulnerabilities of governments in Latin America and the Caribbean. This highlights the pressing need to enhance cybersecurity from multiple perspectives, including strengthening employees' cybersecurity skills. More than two-thirds of cyberattacks on organizations result from employee negligence, with common attacks such as phishing emails tricking users into clicking malicious links, opening virus-infected files, or sharing sensitive information. A meta-analysis of phishing experiments found that an average of 21 percent of employees fall victim to these attacks (Sommestad and Karlzén, 2019). Therefore, it is crucial for public officials to develop the necessary awareness to prevent such attacks.

However, on-the-job training is difficult to implement. Moreover, there is no clear evidence on the most effective strategy to improve employees' learning of general skills. In the public sector, the most common approach to promoting good cybersecurity practices is mandatory training. The literature suggests that individuals primarily invest in education and training to enhance their productivity and earning potential (Becker, 1962) and that they learn more when training aligns with their

interests (Busso et al., 2023). This creates a misalignment of incentives: institutions need employees to act as the first line of defense, but employees rarely see cybersecurity vigilance as essential to their career growth. Another approach occasionally used is phishing campaigns, where employees are intentionally exposed to simulated emails that mimic the tactics of cybercriminals, often followed by educational messages that teach employees how to identify phishing traps. We refer to this approach as "learning-by-doing." Currently, there is no rigorous evidence on the effectiveness of this approach. This study seeks to fill that gap by addressing the central question: What is the most effective strategy to encourage good cybersecurity behavior among public officials?

This study conducted a randomized controlled trial to assess the effectiveness of two strategies for training cybersecurity skills among 1,918 public officials in Argentina: online training and learning-by-doing. Participants were randomly assigned to one of three groups: (i) a group that received promotions for two online cybersecurity courses, (ii) a group that received simulated phishing emails with feedback on their responses, and (iii) a control group with no intervention. The study spanned approximately two months. To measure the impact of the interventions, all participants, regardless of group assignment, received phishing assessment emails both before and after the intervention period. The primary outcomes measured were the proportion of officials who opened the assessment emails, clicked on links within those emails, and reported the phishing attempts. The results suggest that the learning-by-doing approach—sending simulated phishing emails followed by feedback—led to significant improvements in phishing email detection. Participants in this group were less likely to open phishing emails and more likely to report them compared to the control group. Additionally, this approach was particularly effective among permanent public officials compared to contractors, and it showed a stronger impact among female employees.

This study addresses the evidence gap by examining how to effectively promote good cybersecurity behavior in public institutions, where empirical evidence is currently scarce. Existing studies primarily focus on the risks of phishing, revealing that a significant percentage of employees fall victim to these attacks (Sommestad and Karlzén, 2019; Mihelič et al., 2019; Baillon et al., 2019). Various prevention mechanisms have been explored in the literature. Baillon et al. (2019) found that providing information about phishing and allowing public servants to experience phishing attempts in a secure environment reduces the proportion of individuals falling victim to these traps. In contrast, Kim et al. (2018) suggested that threatening employees with disciplinary actions, particularly higher-ranking individuals, is the most effective way to reduce the likelihood of falling for phishing attacks. Jampen et al. (2020), in their

literature review, emphasized the success of phishing training as a preventive measure. Despite these findings, the most effective method for promoting cybersecurity behavior, particularly in the public sector, remains unclear. To address this gap, this research compares two commonly used educational approaches: learning-by-doing and online training. The novelty of this study lies in its experimental comparison of these two approaches with public servants, as well as its analysis of heterogeneous responses based on occupation, institution, and demographic characteristics.

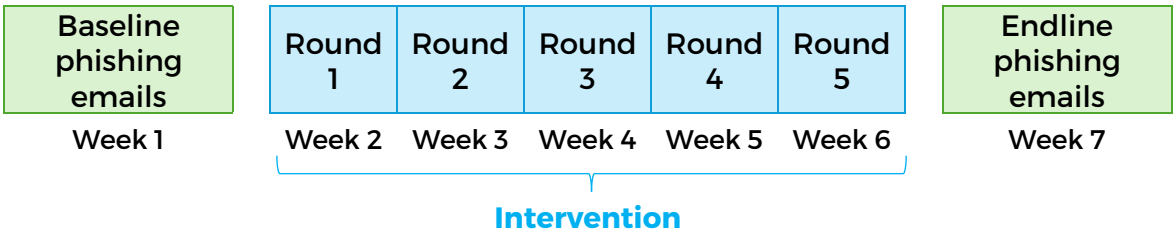
Moreover, this study provides evidence on alternative approaches to promoting general in the workplace. Specifically, it demonstrates the relative effectiveness of the learning-by-doing approach compared to online training. This research adds to the empirical literature by comparing these two learning models in the context of adult education. To the best of our knowledge, only comparative theoretical evidence exists on these models, suggesting that human capital depreciation may be lower under the learning-by-doing model than with online training, as depreciation depends on the intensity of knowledge use (Killingsworth, 1982). In contrast, in early childhood education, both models have been found effective, with stronger evidence supporting the learning-by-doing approach. Araya et al. (2019) demonstrated the effectiveness of learning-by-doing and gamification among primary school students.

The remainder of this paper is organized as follows: Section 2 describes the experimental design, while Section 3 outlines the data and methodology. Section 4 presents a summary of the results, and Section 5 concludes with the main findings and policy implications.

2. Experimental Design

To evaluate the most effective approach to training civil servants in cybersecurity skills, we partnered with the Cabinet of Ministers in Argentina to conduct an experiment between September and October 2023. The experiment was conducted in seven stages (see Figure 1), which included baseline data collection, five rounds of interventions, and endline data collection.

Figure 1. Timeline of the Experiment



Source: Authors' elaboration
Notes: While this timeline reflects the original plan, the actual intervention underwent slight adjustments due to unforeseen challenges during implementation. However, the intervention was successfully completed within an 8-week period

A total of 1,918 active public servants were randomly assigned to one of three groups:

i. **Invitation to online training (639):** Participants in this group received five emails throughout the intervention period, with one email dispatched per round. These emails encouraged them to enroll in a cybersecurity training course, accessible via a direct link provided in each message. The training program included two courses, each lasting 19 to 20 minutes. The first course focused on key strategies for identifying and avoiding phishing attacks, similar to the feedback provided in the learning-by-doing arm. The second course was more interactive, using a game format to cover topics such as password management, phishing and malware protection, and two-factor authentication. Public servants could choose to take either one or both courses.

   To encourage enrollment, the Cabinet of Ministers informed participants that they would receive emails from a reputable cybersecurity company inviting them to participate in a course, ensuring these messages were recognized as legitimate. Additionally, the invitation emails employed several strategies to boost participation. These emails highlighted the importance of cybersecurity training for both personal and workplace protection, emphasized the shared responsibility for maintaining cybersecurity, and underscored the legal obligation to act responsibly. A sense of urgency was created by suggesting that the course had limited availability. However, it is important to note that enrollment in the course remained voluntary.

ii. **Learning-by-doing (639):** Participants completed five rounds of training designed to teach strategies for identifying and avoiding phishing emails. Each round consisted of two emails: first, a simulated phishing email that

mimicked a common tactic used by scammers (see Figure B1), and second, a follow-up email sent within two days to provide performance feedback (see Figure B2). The feedback informed participants whether they had successfully identified and avoided the phishing attempt. Each follow-up email also explained the specific tactic used in the phishing simulation and offered tips for recognizing similar threats in the future. Each round focused on a different phishing tactic: (i) recognizing requests for personal information and urgent appeals; (ii) identifying suspicious links in unexpected emails; (iii) avoiding unsolicited attachments; (iv) spotting spelling errors and scrutinizing emails from unknown senders; (v) combining previously taught techniques to reinforce phishing detection skills.

iii. **Control (640):** Participants in this group did not receive any interventions throughout the experiment.

Randomization was based on a list of public servants provided by the Cabinet of Ministers, which included key demographic characteristics: gender, age, and type of contract. Table 1 presents the balance of observations across the groups. The sample comprises 52 percent women and 48 percent men, with the majority being contractual workers (88 percent) and the remainder permanent workers (12 percent). The age distribution shows that most public servants are younger than 45, with 33 percent between the ages of 18 and 35, and 32 percent between 36 and 45. The proportion of public employees over 60 years old is relatively low, averaging 8.6 percent, with slight variations across groups (8.8 percent in online training and 8.5 percent in both the control and learning-by-doing groups), none of which are statistically significant. These demographic statistics remain consistent across the intervention groups, indicating that the randomization process effectively balanced the demographic characteristics of participants.

To implement the experiment, the research team contracted a platform specialized in cybersecurity capacity building. The selection of courses and the design of the simulated phishing emails were determined through a joint effort between the researchers and the counterpart. All intervention emails were dispatched using this platform, with emails encouraging course enrollment originating from a specific email address. To ensure participant trust and engagement, the Cabinet of Ministers sent an announcement at the beginning of the intervention. This announcement informed participants in the online training arm that they would have the opportunity to undertake online cybersecurity training and instructed them to trust the specific email address from which the training invitations would be sent.

Table 1. Balance of Demographic Characteristics

	Control	Online training		Learning-by-doing	
Variable	Mean	Mean	P-value	Mean	P-value
Sex					
Female	52.0	52.1	0.977	52.0	0.979
Male	48.0	47.9	0.977	48.0	0.979
Type of contract					
Permanent	11.9	12.1	0.931	11.9	1.000
Contractual	88.1	87.9	0.931	88.1	1.000
Age group					
18 to 35	33.1	33.1	1.000	32.9	0.953
36 to 45	32.1	31.8	0.905	32.0	0.952
46 to 59	26.3	26.3	1.000	26.6	0.899
60+	8.5	8.8	0.842	8.5	1.000
N	640	639		639	

Source: Authors' elaboration.
Note: This table presents the mean values for each demographic characteristic within the treatment groups. P-values are the results of t-tests comparing the means between the control group and each treatment group.

3. Data and Methodology of Estimations

Data for the evaluation were collected from two sources. The first source was the cybersecurity training platform, which tracked individualized participant actions in response to the emails. Through this platform, we were able to monitor two of the main outcome variables: email open rates and click rates on links within the emails. To establish a baseline and assess the effectiveness of the interventions, two phishing emails were sent to all participants before the intervention, followed by three phishing emails after the intervention.

The impact of these outcomes may vary depending on the scenario. For example, email open rates could decline if recipients become better at recognizing phishing attempts from the subject line or preview text. However, if public servants follow a policy of opening all emails as part of their due diligence, we may not observe significant changes in open rates. In contrast, the number of clicks on phishing links is a more reliable indicator of the effectiveness of anti-phishing interventions. As participants become more adept at recognizing phishing attempts, we expect a decrease in the number of clicks on these deceptive links. This metric not only offers a clearer measure of intervention success but also sheds light on the sophistication of phishing emails. More convincing phishing attempts may initially result in higher engagement and,

consequently, more clicks, indicating areas where additional training or safeguards are needed.

Descriptive statistics for these outcomes are presented in Table 2. The average open rate for the baseline evaluation emails ranged from 39.5 percent to 45.4 percent, while 10.3 to 11.4 percent of participants clicked on links within those emails. Given the high click rates, it can be inferred that these emails were relatively difficult to detect. The endline evaluation emails were assigned three levels of difficulty, denoted as levels 1, 2, and 3, with level 1 representing the easiest to detect and level 3 the most challenging. The descriptive statistics reflect these difficulty levels. The click rate for level 3 was higher (ranging from 7 to 10.2 percent) than for level 1 (ranging from 0.3 to 0.6 percent).

The cybersecurity capacity-building platform also provided data on treatment take-up rates. For the online training arm, take-up rates were measured by the proportion of participants who either started or completed the course. In the learning-by-doing arm, a proxy measure was used to assess take-up: the percentage of participants who opened any of the feedback emails they received. This proxy served as an indicator of engagement with the learning-by-doing intervention, as opening the feedback emails suggested active participation in the learning process and review of the information provided. Table A1 in Annex A presents descriptive statistics for the take-up rates. Overall, take-up rates for the online training were low: only 4.1 percent of participants in that treatment arm started the course, and just 3 percent completed it. Male, older, and contractual workers were more likely to complete the course. The learning-by-doing arm, while still showing low engagement, fared better: 28 percent of participants opened at least one feedback email. Notably, this arm showed no significant demographic differences in email opening rates.

The second source of information consisted of administrative data collected by the cybersecurity team of the Cabinet of Ministers. Throughout the two-month study period, the team recorded individualized weekly data on participants who reported phishing emails using the Cabinet's established reporting channel. This channel required participants to submit phishing reports to a specific email address designated by the cybersecurity team. Due to the weekly availability of the data, the reporting rates for baseline emails 1 and 2, which were sent in different weeks, can be observed separately. However, all three endline evaluation emails were sent in the same week, so phishing reports for these emails are pooled. Table 2 presents descriptive statistics for this indicator. The data indicate that, on average, one in ten public servants reported phishing emails at both baseline and endline, suggesting a consistent reporting rate throughout the study.

To evaluate the effectiveness of the different cybersecurity training interventions, we employ a combination of ordinary least squares (OLS) regressions, instrumental variable (IV) regressions, and difference-in-differences (DD) models. Our primary analysis relies on OLS regressions to estimate the intention to treat effect of the online training and learning-by-doing interventions on the key outcomes of interest, namely open email rates, click rates, and phishing email reporting rates. The regression model under this specification is:

$$Y_{ip} = \beta\ OT_i + \gamma\ LD_i + \mu X_i + \varepsilon_{ip} \quad \forall p \in \{1,2,3\} \qquad (1)$$

where Y_{ip} is the outcome variable for participant i in endline evaluation phishing email p, OT_i and LD_i are dummy variables that take the value of 1 if the participant was assigned to the online training (OT) or the learning-by-doing (LD) arm, respectively, and zero otherwise; X_i is a vector of individual characteristics, including gender, age, type of contract, and baseline knowledge as measured by baseline outcomes; and ε_{ip} is the error term. Separate regressions are run for each evaluation email, but joint results are presented in the Appendix. Note that in all regression tables report the results for β and γ, along with the standardized effect size, calculated using Cohen's d, defined as:

$$Effect\ size = \frac{\bar{Y}_{treatment} - \bar{Y}_{control}}{SD_{pooled}} \qquad (2)$$

where $\bar{Y}_{treatment}$ is the mean of the outcome in the treatment group (whether online training or learning-by-doing), $\bar{Y}_{control}$ is the mean of the control group, and SD_{pooled} is the standard deviation across the two groups.

Table 2. Descriptive Statistics of Outcome Variables

	Open Rate (%)			Click Rate (%)			Phishing Report Rate (%)		
Type of email	Control	Online training	Learning-by-doing	Control	Online training	Learning-by-doing	Control	Online training	Learning-by-doing
Baseline									
E-mail A	39.5	43.7	40.5	10.3	11.4	10.5	11.1	10.5	7.7
E-mail B	43.1	45.4	42.1	10.6	11.3	11.0	10.2	10.0	10.2
Endline									
E-mail 1	60.6	62.6	57.0	0.6	0.3	0.6			
E-mail 2	13.6	17.8	13.3	3.9	4.5	1.9	11.2	10.6	11.3
E-mail 3	62.5	64.6	55.9	9.1	10.2	7.0			

Source: Authors' elaboration

Notes: This table presents summary statistics of the main outcome variables, organized by columns. The data on opened emails and clicks originate from the experiment's platform, while reported phishing incidents are sourced from administrative records provided by the Cabinet of Ministers in Argentina. The experiment entailed sending two evaluation phishing emails before the intervention and three emails at the end. Each row presents statistics for the corresponding email. Endline emails were dispatched in the same week, and phishing reports were collected weekly. Therefore, individualized data per email are unavailable for this variable.

OLS regressions on experimental data that ignore non-compliance can result in attenuated treatment effects (Angrist and Pischke, 2009). To address potential self-selection bias in treatment take-up, we complement OLS regressions with instrumental variable (IV) regressions. In the IV approach, we use treatment assignment as an instrument for actual take-up rates, allowing us to estimate the local average treatment effect (LATE). This focuses on the subpopulation of compliers—those who engaged with the offered trainings—thereby mitigating concerns about actual exposure to the treatment. The IV results can be interpreted as an upper bound on the treatment's potential effectiveness, representing the effect if all aspects of the intervention proceed as intended. However, it's important to note that in real-world scenarios, compelling employees to open emails, let alone read them, is challenging. Training is not always mandatory, and even when it is, it doesn't guarantee effective learning. Consequently, the IV results for each treatment arm should be understood as the maximum potential impact of these trainings under ideal conditions.

Finally, as a robustness check and to explore the impact of the interventions over time, we estimate difference-in-differences (DD) models[1] that leverage the variation in outcomes between the baseline and endline periods. These models include a time dummy variable that takes the value of 1 for endline emails and 0 for baseline emails, along with an interaction term between this dummy and the treatment dummies. The DD approach controls for time-invariant differences between the treatment and control groups, as well as common time trends that may affect all groups equally.

In this analysis, we cluster standard errors at the department level to account for potential correlations in outcomes among individuals within the same team.

4. Results

We first estimate the results for email open rates. Table 3 shows that the online training and learning-by-doing interventions had varying effects on open rates for the three endline phishing evaluation emails, which differed in difficulty. The online training treatment had a positive and statistically significant effect on the open rate for the moderately difficult level 2 email but showed no effect on the easiest (level 1) or most difficult (level 3) emails. In contrast, the learning-by-doing treatment reduced the open rate for the most difficult email (level 3) by 6 percentage points in the OLS estimation. When adjusting for take-up rates using the IV approach, the decrease in open rates for

[1] The Difference-in-Differences model is specified as: $Y_{it} = \beta_1 \, OT_i + \beta_2 \, (OT_i \times After_t) + \gamma_1 LD_i + \gamma_2 (LD_i \times After_t) + \mu X_i + \rho_t + \varepsilon_{it}$, where $After_t$ takes the value of 1 at endline and 0 at baseline, and ρ_t is a vector of dummies controlling for the type of email round.

the level 3 email becomes even more pronounced, reaching 23.7 percentage points. The IV results can be interpreted as an upper bound on the potential effectiveness of the treatment arms, particularly training. While it is difficult to compel employees to open or read emails, training can sometimes be made mandatory. Thus, the IV result for the training arm can be understood as the maximum potential impact of mandatory training. The learning-by-doing treatment, however, did not significantly affect open rates for the other endline emails. Table A2 shows that these effects hold with the difference-in-differences estimation.

The interpretation of these results is complex, as opening a phishing email does not necessarily indicate cyber-risky behavior. While opening a phishing email could pose a potential threat, it is also possible that individuals open emails because they are diligent in managing their inbox, without exposing the organization to risk. Conversely, individuals who do not open emails may be more distracted and, as a result, could be more susceptible to phishing when they do engage with an email. Therefore, while these results offer insights into the effects of the interventions, examining click rates alongside open rates could provide a more comprehensive understanding of the interventions' impact on cybersecurity risks.

Turning to the results on click rates, the analysis indicates that online training generally did not have statistically significant effects on click rates within phishing emails, as shown in Table 4, while learning-by-doing had moderate effects. The online training treatment did not significantly impact click rates for any of the three endline emails, as evidenced by the small and statistically insignificant coefficients in both the OLS and IV regressions. This suggests that merely offering online training, despite its potential to raise awareness, may not have been sufficient to induce changes in click behavior. In contrast, the learning-by-doing treatment showed a negative and statistically significant effect on the click rate for the second endline email (moderate difficulty) in both the OLS and IV estimates. This finding suggests that the learning-by-doing approach may be somewhat effective in helping participants identify and avoid clicking on links in moderately difficult phishing emails. However, the overall effectiveness of the learning-by-doing intervention in reducing click rates appears limited, possibly due to the already low click rates among participants.

Table 3. Effects of Online Training and Learning-By-Doing on Email Open Rate

	Endline 1		Endline 2		Endline 3	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Online training						
Beta	0.019	0.646	0.041 **	1.391 *	0.022	0.755
SE	(0.027)	(0.919)	(0.020)	(0.728)	(0.027)	(0.908)
Effect size	0.016	0.016	0.046	0.044	0.019	0.019
First stage F-test		19.7		19.6		19.7
Learning-by-doing						
Beta	-0.036	-0.130	-0.004	-0.013	-0.066 **	-0.237 **
SE	(0.027)	(0.100)	(0.019)	(0.069)	(0.027)	(0.101)
Effect size	-0.030	-0.030	-0.004	-0.004	-0.055	-0.054
First stage F-test		244.9		245.2		244.9
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$ (p-value)	0.043	0.375	0.026	0.045	0.001	0.053
Control mean	0.606	0.606	0.136	0.136	0.625	0.625
Observations	1914	1914	1908	1908	1914	1914

Source: Authors' elaboration.

Notes: Each column presents results from a separate regression where the dependent variable is the email open rate for each of the endline phishing evaluation emails. The outcome variable takes the value of 1 if the experiment participant opened the respective endline phishing email and 0 otherwise. Columns 1 and 2 use the outcome of the first endline email, columns 3 and 4 for the second email, and columns 5 and 6 for the third email. Odd-numbered columns (1, 3, and 5) report estimates from OLS regressions, while even-numbered columns (2, 4, and 6) report estimates from instrumental variable (IV) regressions where treatment assignment instruments the take-up rates of each treatment. All regressions control for participant gender, age, and job function. "OT" denotes the Online Training treatment, and "LD" denotes the Learning by Doing treatment. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are shown in the third row of each set of coefficients. ***p<0.01, **p<0.05, *p<0.1.

Table 4. Effects of Online Training and Learning-By-Doing on Email Click Rate

	Endline 1		Endline 2		Endline 3	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Online training						
Beta	-0.003	-0.105	0.006	0.204	0.011	0.366
SE	(0.004)	(0.131)	(0.011)	(0.382)	(0.017)	(0.562)
Effect size	-0.019	-0.018	0.012	0.012	0.015	0.015
First stage F-test		19.7		19.6		19.7
Learning-by-doing						
Beta	-0.000	-0.000	-0.021**	-0.074**	-0.020	-0.074
SE	(0.004)	(0.016)	(0.009)	(0.035)	(0.015)	(0.056)
Effect size	-0.000	-0.000	-0.050	-0.049	-0.031	-0.030
First stage F-test		244.9		245.2		244.9
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$						
(p-value)	0.417	0.392	0.007	0.446	0.046	0.412
Control mean	0.006	0.006	0.039	0.039	0.091	0.091
Observations	1914	1914	1908	1908	1914	1914

Source: Authors' elaboration.

Notes: Each column presents results from a separate regression, with the dependent variable being the email click rate for each of the endline phishing evaluation emails. The outcome variable is coded as 1 if the participant clicked on a link in the respective endline phishing email and 0 otherwise. Columns 1 and 2 correspond to the first endline email, columns 3 and 4 to the second email, and columns 5 and 6 to the third email. Odd-numbered columns (1, 3, and 5) report estimates from OLS regressions, while even-numbered columns (2, 4, and 6) report estimates from instrumental variable (IV) regressions, where treatment assignment instruments the take-up rates of each treatment. All regressions control for participant gender, age, and job function. "OT" refers to the Online Training treatment, and "LD" refers to the learning-by-doing treatment. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are shown in the third row of each set of coefficients. ***p<0.01, **p<0.05, *p<0.1.

Moving on to the results on phishing email reporting, as shown in Table 5, neither the online training nor the learning-by-doing interventions had a statistically significant impact on the likelihood of reporting phishing emails. The OLS estimates in column 1 show small, negative, and statistically insignificant coefficients for both treatments, indicating that participants in the treatment groups were no more likely to report phishing emails than those in the control group. These findings are consistent with the overall pattern of limited effects observed in the interventions, as previously noted with the click rates. The IV and DD estimates further support these results, failing to detect any significant effects of the interventions on reporting behavior. The absence of significant results suggests that while the interventions may have influenced participants' awareness or email-clicking behavior, they were not effective in

encouraging participants to actively report phishing emails, despite the potential benefits to the institution's cybersecurity.

Table 5. Effects of Online Training and Learning-By-Doing on Phishing Report Rate

	OLS (1)	IV (2)	DD (3)
Online training			
Beta	-0.007	-0.219	-0.002
SE	(0.017)	(0.583)	(0.017)
Effect size	-0.009	-0.009	-0.002
First stage F-test		19.7	
Learning-by-doing			
Beta	-0.000	-0.001	0.017
SE	(0.018)	(0.063)	(0.017)
Effect size	-0.000	-0.000	0.013
First stage F-test		244.9	
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$ (p-value)	0.713	0.693	0.253
Control mean	0.112	0.112	0.112
Observations	1914	1914	5735

Source: Authors' elaboration.
Notes: Each column presents results from a separate regression model where the dependent variable is a binary indicator of whether the participant reported the phishing emails during the evaluation week. Since the three evaluation emails were sent within the same week, there is only one endline observation per participant. Column 1 reports OLS estimates, column 2 reports IV estimates using treatment assignment as an instrument for treatment take-up, and column 3 reports DD estimates from the interaction term between the intervention dummy variable and a binary variable that equals 1 for endline emails and 0 for baseline emails. All models control for participant gender, age, and job function. "OT" denotes Online Training and "LD" denotes Learning by Doing. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are shown in the third row of each set of coefficients. ***p<0.01, **p<0.05, *p<0.1.

The relatively low baseline reporting rate, combined with the absence of significant effects, suggests that factors beyond awareness and training may have played a more decisive role in whether employees reported phishing attempts. One plausible explanation is that the cost of reporting was too high. Public servants were required to send an email to the cybersecurity team to report phishing attempts, which may have posed a barrier if they had to search for the correct email address or were uncertain about where to send the report. This extra effort could have discouraged participants from reporting phishing emails, even if they were trained to identify them. However, without further data, it is challenging to pinpoint the precise mechanisms behind the low reporting rate and the limited effectiveness of these interventions in encouraging reporting behavior.

Heterogeneous effects of the interventions based on employees' demographic characteristics were also examined. Table 6 presents OLS estimations by type of contract, indicating that the learning-by-doing intervention was more effective at increasing phishing email reporting rates among permanent public servants compared to contractual staff. This may be due to permanent employees' stronger alignment with the organization's cybersecurity goals and their longer-term commitment to the institution. No significant differential effects of learning-by-doing were found for email open rates or click rates. Similarly, no significant heterogeneous effects of the online training intervention were observed based on contract type for any of the outcomes considered.

The effectiveness of the cybersecurity training interventions was also examined for potential variation by gender. Table 7 presents OLS regression results, showing that the online training intervention did not have significant gender-differentiated effects on any of the outcome variables. In contrast, a significant positive effect of the learning-by-doing intervention was observed on the likelihood of reporting phishing emails among female employees compared to their male counterparts. Specifically, the interaction term between the learning-by-doing treatment and gender is positive and statistically significant (14.2 percentage points), suggesting that female employees were more likely to report phishing emails after undergoing the learning-by-doing training. This gender difference may be partially attributed to higher engagement among women, as reflected in their higher open rates for feedback emails (see Table A1). However, no significant gender differences were found for email open rates and click rates within the learning-by-doing intervention.

As an additional robustness check, pooled regressions were run without differentiating between the difficulty levels of the endline emails. The results, displayed in Table A4, align with the overall findings previously discussed: online training had no significant impact on any of the three outcome variables, while the learning-by-doing intervention led to a reduction in both email open rates and click rates on links within those emails. Neither intervention, however, had any effect on phishing email reporting.

Several hypotheses could explain the lack of significant effects from the online training intervention and the effects observed in the learning-by-doing arm. First, low take-up rates for both interventions, particularly the online training, which was voluntary. Despite efforts to encourage participation, including multiple reminder emails and emphasizing the importance of cybersecurity, only a small percentage of participants enrolled in the courses. The voluntary nature of the intervention likely led to self-

selection bias, where only those with a pre-existing interest in cybersecurity participated, limiting the potential for widespread behavior change.

Another possible explanation lies in the design of the interventions themselves. The online training may not have been interactive or engaging enough to lead to substantial behavior change, as participants were merely exposed to information without actively applying it. In contrast, the learning-by-doing approach showed some limited success in phishing email reporting, but its overall impact was likely constrained by the already low baseline click rates, leaving little room for improvement. The reporting mechanism, which required public servants to actively email the cybersecurity team, may have also deterred individuals from reporting phishing emails, even if they had been trained to recognize them. Finally, while the experiment may seem underpowered, post-estimation power calculations indicate a high level of statistical power given the sample size and observed effect sizes. This suggests that the issue may not be related to statistical power (see Figure A1).

Table 6. OLS Heterogeneous Effects, by Type of Contract

		Opened emails			Clicks on emails			Reported emails
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Online traning	Beta	0.070	0.023	0.047	-0.005	-0.007	-0.014	0.077
	SE	(0.060)	(0.055)	(0.060)	(0.026)	(0.028)	(0.032)	(0.051)
	Effect size	0.015	0.006	0.010	-0.003	-0.003	-0.006	0.020
Learning-by-doing	Beta	0.088	0.064	0.091	0.026	0.018	0.024	0.142 ***
	SE	(0.058)	(0.053)	(0.059)	(0.024)	(0.026)	(0.030)	(0.049)
	Effect size	0.020	0.016	0.021	0.015	0.009	0.010	0.038
Endline		1	2	3	1	2	3	1,2,3
Observations		5735	5729	5735	5735	5729	5735	5735

Source: Authors' elaboration.

Notes: Each column presents results from a separate OLS regression model, where the dependent variable is the outcome mentioned in the column header. The independent variables in each model include the treatment variable, the heterogeneous variable (type of contract), and an interaction term between the type of contract and the treatment variable. The estimates shown in the table correspond to the coefficient of the interaction term, which captures the differential impact of the treatment on the outcome for permanent employees compared to temporary participants. The type of contract is a binary variable that takes the value of 1 for permanent employees and 0 for temporary participants. Since the three evaluation emails were sent within the same week, there is only one endline observation per participant. All models control for participant gender, age, and job function. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are shown in the third row of each set of coefficients. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. OLS Heterogeneous Effects, by Gender

		Opened emails			Clicks on emails			Reported emails
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Online traning	Beta	-0.044	-0.035	-0.021	-0.009	-0.009	-0.007	0.011
	SE	(0.038)	(0.033)	(0.039)	(0.018)	(0.020)	(0.021)	(0.034)
	Effect size	-0.015	-0.014	-0.007	-0.007	-0.006	-0.005	0.004
Learning-by-doing	Beta	0.001	0.019	-0.001	0.029	0.029	0.023	-0.062 *
	SE	(0.038)	(0.032)	(0.038)	(0.018)	(0.019)	(0.021)	(0.034)
	Effect size	0.000	0.008	-0.000	0.021	0.020	0.014	-0.024
Endline		1	2	3	1	2	3	1,2,3
Observations		5735	5729	5735	5735	5729	5735	5735

Source: Authors' elaboration.

Notes: Each column presents results from a separate OLS regression model, where the dependent variable is the outcome mentioned in the column header. The independent variables in each model include the treatment variable, the heterogeneous variable (gender), and an interaction term between the participant's gender and the treatment variable. The estimates shown in the table correspond to the coefficient of the interaction term, which captures the differential impact of the treatment on the outcome for male compared to female employees. Gender is a binary variable that takes the value of 1 for male employees and 0 for female employees. Since the three evaluation emails were sent within the same week, there is only one endline observation per participant. All models control for participant gender, age, and job function. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are shown in the third row of each set of coefficients. *** p<0.01, ** p<0.05, * p<0.1.

5. Conclusions and Policy Implications

This study investigates the effectiveness of two cybersecurity training interventions - online training and learning-by-doing - in improving cybersecurity skills and behaviors among public servants in Argentina. Through a randomized controlled trial involving 1,918 participants, we assessed the impact of these interventions on email opening, clicking, and reporting rates for simulated phishing emails of varying difficulty levels.

Our findings suggest that the online training intervention had limited effects on participants' cybersecurity behaviors. While it increased the likelihood of opening moderately difficult phishing emails, it did not significantly impact click rates or reporting behavior. In contrast, the learning-by-doing intervention, which involved sending participants simulated phishing emails and providing feedback on their responses, showed more promising results. This approach led to a reduction in click rates for moderately difficult phishing emails and had a positive effect on the likelihood of reporting phishing attempts, particularly among permanent employees and female participants.

These results have significant policy implications for public institutions aiming to improve cybersecurity behaviors. First, our findings indicate that practical and interactive interventions, such as learning-by-doing through phishing simulations, are slightly more effective in promoting cybersecure behavior among public employees compared to traditional online training. The hands-on experience provided by the learning-by-doing approach allows employees to practice identifying and responding to simulated phishing attempts in a safe environment, which may lead to better outcomes in terms of reduced click rates and increased reporting of phishing attempts. However, it is important to note that the limited effectiveness of online training in the context of this experiment may be attributed to its non-mandatory nature. Given the lack of incentives for employees to acquire these general skills, public institutions should consider making cybersecurity training compulsory for all employees to ensure higher engagement and participation rates.

Second, the low phishing reporting rates observed in this study suggest that public institutions need to do more to encourage employees to report phishing attempts. One way to address this issue is by simplifying the reporting process, for example, by integrating alert buttons into the employee's email platform. This would make it easier and more convenient for employees to report suspicious emails without having to navigate to a separate system or compose an email to the cybersecurity team. Additionally, public institutions should emphasize the importance of reporting

phishing attempts for the organization's overall cybersecurity. This can be achieved through regular communication campaigns, posters, and other awareness-raising activities that highlight the critical role employees play in protecting the institution from cyber threats.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. "Certification of Training and Training Outcomes." European Economic Review 44, no. 4–6 (2000): 917–927.

Angrist, Joshua D., and Jörn-Steffen Pischke. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press, 2009.

Baillon, Aurélien, Joep P. De Bruin, Ali Emirmahmutoglu, Eva Van De Veer, and Bram Van Dijk. "Informing, Simulating Experience, or Both: A Field Experiment on Phishing Risks." PLOS ONE 14, no. 12 (2019): e0224216.

Becker, Gary S. "Investment in Human Capital: A Theoretical Analysis." Journal of Political Economy 70, no. 5, Part 2 (1962): 9–49.

Busso, Matias, Kyung Park, and Nestor Irazoque. "The Effectiveness of Management Training Programs: A Meta-Analytic Review." 2023.

Cybint. "15 Alarming Cybersecurity Facts and Stats." 2022. https://www.cybintsolutions.com/cyber-security-facts-stats/.

Cybercrime Magazine. "Cybercrime Damages $6 Trillion by 2021." 2017. https://cybersecurityventures.com/annual-cybercrime-report-2017/.

Jampen, David, Gur Gür, Tobias Sutter, and Bernhard Tellenbach. "Don't Click: Towards an Effective Anti-Phishing Training. A Comparative Literature Review." Human-centric Computing and Information Sciences 10, no. 1 (2020): 1–41.

Killingsworth, Mark R. ""Learning by Doing" and "Investment in Training": A Synthesis of Two "Rival" Models of the Life Cycle." Review of Economic Studies 49, no. 2 (1982): 263–271.

Kim, Bongjun, Deokjin Y. Lee, and Bonkyu Kim. "Deterrent Effects of Punishment and Training on Insider Security Threats: A Field Experiment on Phishing Attacks." Behaviour & Information Technology 39, no. 11 (2020): 1156–1175.

Mihelič, Andrej, Miha Jevšček, Sašo Vrhovec, and Igor Bernik. "Testing the Human Backdoor: Organizational Response to a Phishing Campaign." Journal of Universal Computer Science 25, no. 11 (2019): 1458–1477.

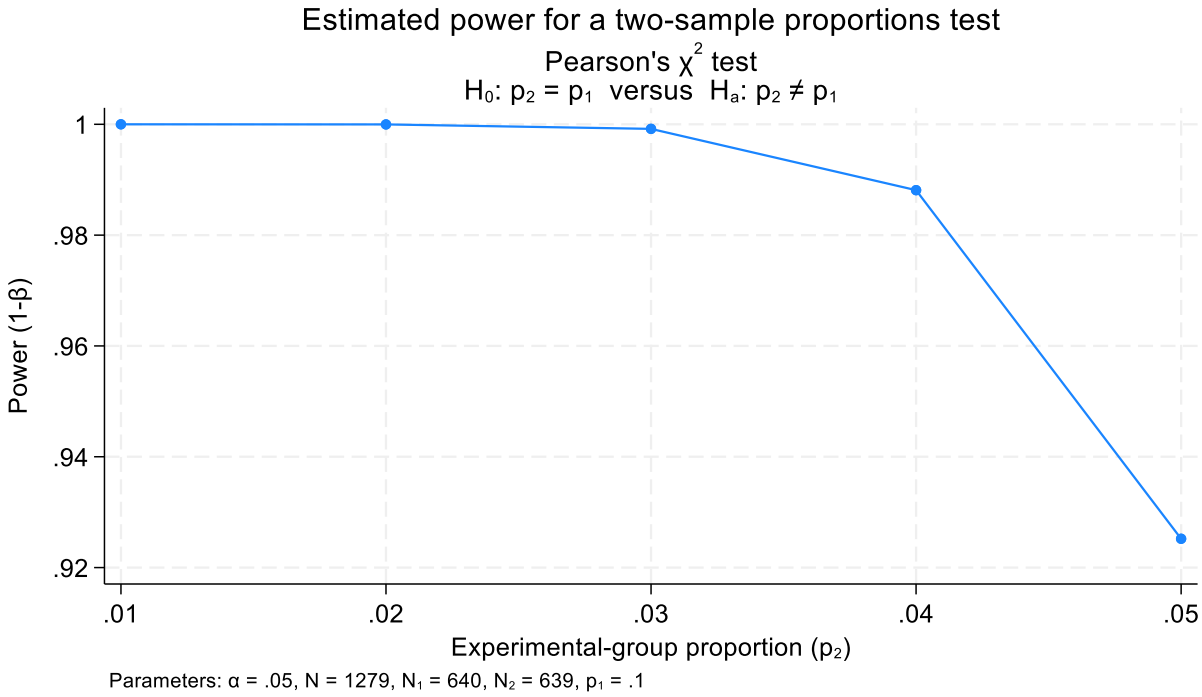
Offerman, Theo. "Hurting Hurts More Than Helping Helps." European Economic Review 46, no. 8 (2002): 1423–1437.

Sommestad, Teodor, and Henrik Karlzén. "A Meta-Analysis of Field Experiments on Phishing Susceptibility." In 2019 APWG Symposium on Electronic Crime Research (eCrime), 1–14, November 2019.

World Economic Forum (WEF). The Global Risks Report 2022: Insight Report. 2022. https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2022.pdf.

Annex A. Additional Tables and Figures

Figure A1. Post-Estimation Power Calculations



Source: Authors' elaboration.
Notes: The graph shows power calculation results based on a baseline click rate of 10 percent for email links, with potential endline click rates shown on the horizontal axis. The control group consists of 640 participants, and the treatment group consists of 639, reflecting the actual sample sizes from the experiment. The significance level (alpha) is set at 0.05.

Table A1. Intervention Take-up Rates, by Demographic Group

	Online training		Learning-by-doing
	Started course	Completed course	Opened feedback email
	4.1	3.0	27.7
Gender			
Female	2.7	1.5	29.5
Male	5.6	4.6	25.7
Age			
18 to 35	1.9	1.4	26.7
36 to 45	4.9	3.4	26.5
46 to 59	4.2	3.0	27.1
60+	8.9	7.1	38.9
Type of contract			
Permanent	2.6	1.3	23.7
Contractual	4.3	3.2	28.3

Source: Authors' elaboration.
Note: The data on course completion and opened emails originate from the experiment's platform.

Table A2. Difference-in-Differences of Online Training and Learning-By-Doing on Email Open Rate

	(1)	(2)	(3)	(4)	(5)
Online Training					
Beta	-0.003	-0.004	-0.013	0.012	-0.011
SE	(0.019)	(0.019)	(0.026)	(0.024)	(0.025)
Effect	-0.001	-0.002	-0.006	0.006	-0.006
Learning-by-doing					
Beta	-0.035 *	-0.035 *	-0.037	-0.002	-0.066 ***
SE	(0.019)	(0.019)	(0.025)	(0.023)	(0.025)
Effect	-0.019	-0.019	-0.019	-0.001	-0.035
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$ (p-value)	0.098	0.111	0.345	0.561	0.031
N	9557	9577	5747	5741	5747
Controls	Yes	No	No	No	No
Fixed effects	No	Yes	Yes	Yes	Yes
Endline type	1,2,3	1,2,3	1	2	3

Source: Authors' elaboration.
Notes: Each column presents results from a separate difference-in-differences regression model where the dependent variable is the email open rate for the endline phishing evaluation emails. The outcome variable takes the value of 1 if the experiment participant opened the respective endline phishing email and 0 otherwise. Column 1 uses the combined outcome of all three endline emails, while columns 2 to 4 report estimates for each individual endline email. The coefficients shown in the table represent the interaction term between the intervention dummy variable and a binary variable that equals 1 for endline emails and 0 for baseline emails. This interaction term captures the treatment effect of the interventions on the change in email open rates from baseline to endline. All regressions control for participant gender, age, and type of contract. Columns 2 to 4 additionally include individual fixed effects to account for time-invariant individual characteristics. "OT" denotes the Online Training treatment, and "LD" denotes the Learning by Doing treatment. Robust standard errors, clustered at the individual level, are displayed in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A3. Difference-in-Differences of Online Training and Learning-By-Doing on Email Click Rate

	(1)	(2)	(3)	(4)	(5)
Online Training					
Beta	-0.002	-0.004	-0.012	-0.002	0.003
SE	(0.014)	(0.014)	(0.013)	(0.015)	(0.020)
Effect	-0.002	-0.003	-0.011	-0.002	0.002
Learning-by-doing					
Beta	-0.016	-0.016	-0.002	-0.023	-0.022
SE	(0.014)	(0.014)	(0.014)	(0.015)	(0.019)
Effect	-0.012	-0.012	-0.002	-0.020	-0.015
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$ (p-value)	0.320	0.382	0.495	0.180	0.194
N	9557	9577	5747	5741	5747
Controls	Yes	No	No	No	No
Fixed effects	No	Yes	Yes	Yes	Yes
Endline type	1,2,3	1,2,3	1	2	3

Source: Authors' elaboration.

Notes: Each column presents results from a separate difference-in-differences regression model where the dependent variable is the email open rate for the endline phishing evaluation emails. The outcome variable takes the value of 1 if the experiment participant opened the respective endline phishing email and 0 otherwise. Column 1 uses the combined outcome of all three endline emails, while columns 2 to 4 report estimates for each individual endline email. The coefficients shown in the table represent the interaction term between the intervention dummy variable and a binary variable that equals 1 for endline emails and 0 for baseline emails. This interaction term captures the treatment effect of the interventions on the change in email open rates from baseline to endline. All regressions control for participant gender, age, and type of contract. Columns 2 to 4 additionally include individual fixed effects to account for time-invariant individual characteristics. "OT" denotes the Online Training treatment, and "LD" denotes the Learning by Doing treatment. Robust standard errors, clustered at the individual level, are displayed in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A4. OLS and IV Results with Pooled Baseline and Endline Indicators

	Opened emails		Clicks on emails		Phishing reported	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Online training						
Beta	0.011	0.256	0.004	0.091	-0.006	-0.136
SE	(0.016)	(0.402)	(0.006)	(0.154)	(0.016)	(0.391)
Effect size	0.009	0.008	0.008	0.008	-0.005	-0.005
First stage F-test		28.2		28.2		
Learning-by-doing						
Beta	-0.035 **	-0.127 **	-0.014 **	-0.050 *	0.006	0.021
SE	(0.017)	(0.061)	(0.007)	(0.026)	(0.015)	(0.055)
Effect size	-0.028	-0.027	-0.026	-0.026	0.005	0.005
First stage F-test		204.6		204.6		
$\hat{\beta}_{OT} = \hat{\beta}_{LD}$ (p-value)	0.713	0.053	0.009	0.356	0.455	0.653
Observations	5709	5709	5709	5709	5709	5709

Source: Authors' elaboration.

Notes: Each column presents results from a separate regression, with the dependent variables indicated in the column headers. Odd-numbered columns (1, 3, and 5) report estimates from OLS regressions, while even-numbered columns (2, 4, and 6) report estimates from instrumental variable (IV) regressions, where treatment assignment is used as an instrument for treatment take-up rates. All regressions control for participant gender, age, and job function. Clustered standard errors at the department level are reported in parentheses. Standardized effect sizes are provided in the third row of each set of coefficients. ***p<0.01, **p<0.05, *p<0.1.

Annex B. Intervention Details

Figure B1. Example of an Email Encouraging Public Servants to Enroll in Online Training

Protege a la Jefatura y a tu hogar aprendiendo sobre riesgos cibernéticos

From: Security Education Platform

Estimado/a,

Los riesgos cibernéticos representan una amenaza tanto para tu entorno laboral como para tu hogar. La protección ante estos riesgos es una responsabilidad que compartimos, respaldada por la ley. Nuestras familias y la institución cuentan con tu apoyo para defendernos de las crecientes amenazas cibernéticas.

Por lo tanto, la Jefatura de Gabinete te invita a tomar un curso en el que aprenderás cuáles son los riesgos asociados al uso de internet, como el phishing, y diferentes estrategias para evitar caer en trampas cibernéticas. Este curso tiene una duración de 39 minutos y estará disponible por tiempo limitado.

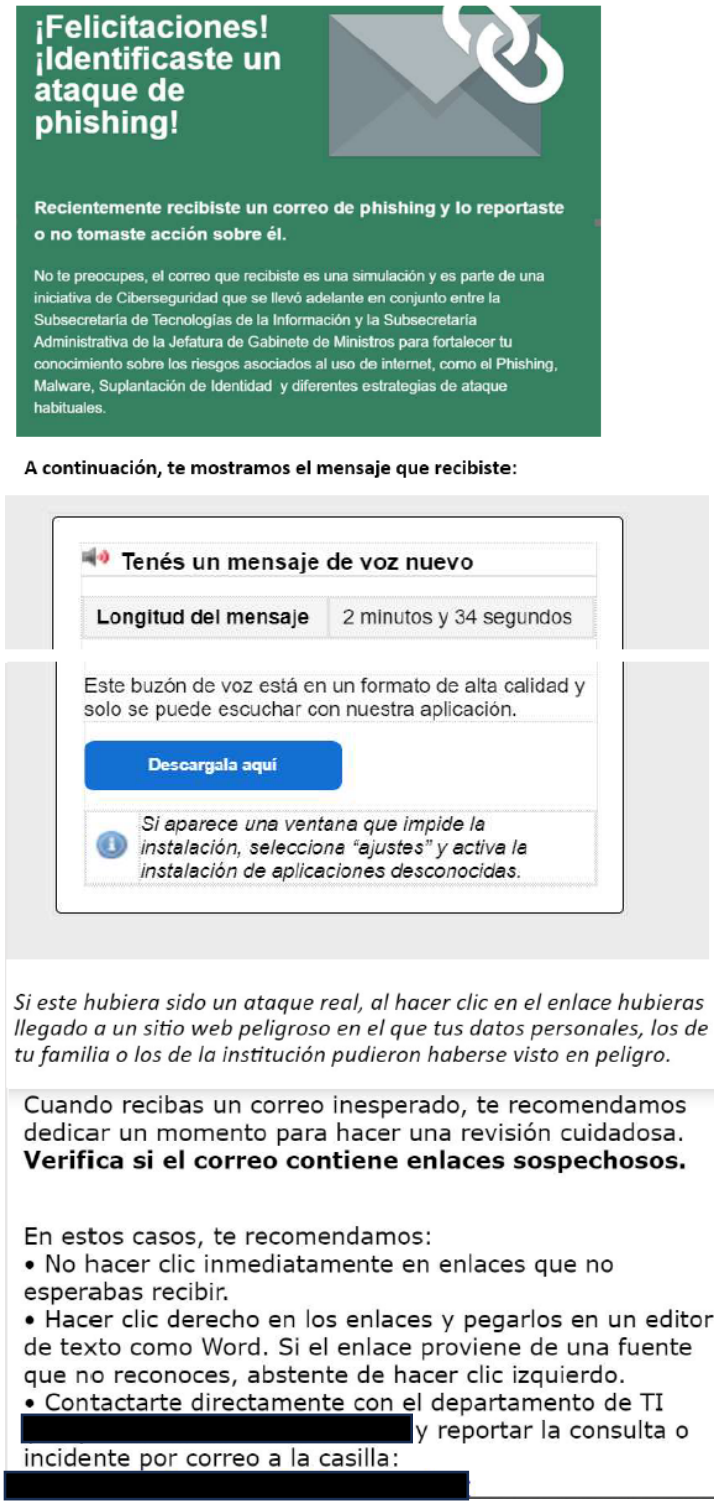
Inscríbete ya en este link

https://iadb.ws01-securityeducation.com/ticketAuth/assignmentTicket

Saludos

Source: Proofpoint Cybersecurity.

Figure B2. Example of Positive Feedback Received During the Intervention Period



Source: Proofpoint Cybersecurity.