Phishing Resistance in 2026: A Ready-to-Deliver Employee Training Program and Handbook

Audience: All employees, contractors, people managers, executives, finance, payroll, HR, IT/help desk, procurement, and accounts payable **Program owner:** Security Awareness / Information Security **Version:** 1.0 — July 2026 **Recommended live delivery:** 60–75 minutes, plus role-based modules and continuous reinforcement **Core behavior: Stop. Think. Verify. Report.**

> ***Manager's opening script***
>
> *Phishing is no longer mainly a badly written email. It can be a flawless message, a text, a QR code, a Teams or Slack chat, or a convincing call or video that appears to come from someone you know. You are not expected to be a forensic analyst, and you will not be blamed for reporting something legitimate. Your job is to pause when a message asks you to cross a trust boundary—open a link, sign in, approve MFA, disclose information, change payment details, or send money—verify the request through a channel you choose, and report it. Fast reporting protects everyone, even after a click.*

Executive brief

This program turns phishing awareness from a once-a-year recognition exercise into a practiced business process. It teaches every employee one response pattern:

1. **Stop** before taking an irreversible or sensitive action.
2. **Think** about the request, not merely the message's appearance.
3. **Verify** high-risk requests through a trusted, independent channel.
4. **Report** suspicious contact immediately, including uncertainty and mistakes.

That pattern matters because the sender name, writing quality, caller ID, face, voice, logo, conversation history, and even the sending account can all be genuine-looking or compromised. In 2025, the FBI's Internet Crime Complaint Center received 191,561 phishing/spoofing complaints. It recorded 24,768 business email compromise (BEC) complaints and $3.047 billion in reported BEC losses; these are complaints reported to IC3, not a census of all victimization (FBI IC3 2025 Annual Report, pp. 6–9). APWG, using a different collection method, counted 971,181 reported phishing attacks in Q1 2026, 13.8% more than in Q4 2025, while warning that changes in email systems can reduce the number of samples forwarded to its collection address (APWG Q1 2026 report, pp. 3–4). These data sets measure different things and should not be added together.

Attackers are also changing channels. Verizon's 2026 DBIR reports that mobile-centered interactive social engineering through texts and calls had a 40% higher success rate than traditional email phishing in its data (Verizon 2026 DBIR release). APWG's contributing data found telephone-based fraud—vishing plus SMS/text phishing—rose 15% from Q4 2025 to Q1 2026 (APWG Q1 2026 report, p. 4). Microsoft observed persistent Teams voice-phishing in a 2025 customer intrusion, in which an attacker impersonated IT support and used legitimate Windows mechanisms as part of the compromise (Microsoft Incident Response, March 2026).

AI changes the cost, speed, scale, personalization, and polish of deception; it does not change the safe decision. The FBI says generative tools can produce official-sounding executive messages and that voice cloning can support payment fraud. IC3 received 22,364 complaints marked AI-related in 2025, with $893.3 million in adjusted reported losses, including more than $30 million reported in AI-related BEC. The FBI explicitly notes that not every BEC attack is AI-enabled (FBI IC3 2025 Annual Report, pp. 38–39). These figures depend on complainants recognizing and reporting an AI nexus, so they establish documented harm, not AI's total prevalence.

Training helps people rehearse safe decisions and creates more detectors, but it is not a firewall. A 2025 randomized field experiment involving more than 19,500 healthcare employees and ten campaigns found no relationship between recency of annual awareness training and simulation failure. Embedded training produced only a two-percentage-point average reduction, more than half of training-page visits ended within ten seconds, and fewer than 24% formally completed the material (UC San Diego/IEEE Symposium on Security and Privacy 2025 paper). An earlier healthcare study likewise found that a mandatory course for repeatedly susceptible employees did not substantially change their click rate relative to others (Gordon et al., *JAMIA*, 2019). NIST cautions that click rates vary with lure difficulty and the recipient's work context; a low rate may mean the exercise was easy, not that people became secure (NIST Phish Scale research).

Accordingly, this program does not promise a "human firewall." It combines:

- short, repeated, role-relevant learning and simulations;
- a visible, no-blame reporting path;
- independent verification and dual-control business processes;
- phishing-resistant MFA, preferably FIDO/WebAuthn;
- email, identity, endpoint, collaboration-platform, and payment controls; and

- practiced containment and recovery.

Learning outcomes

By the end of the core session, participants will be able to:

- recognize phishing across email, voice, video, SMS, QR codes, and collaboration tools;
- inspect the actual sender address, registrable domain, and link destination;
- explain why good grammar, a familiar voice, a real conversation thread, and a valid HTTPS padlock are not proof;
- refuse unexpected MFA prompts and report MFA fatigue;
- use out-of-band verification for every credential, money, payroll, banking, gift-card, sensitive-data, or remote-access request;
- report suspicious messages with the company’s one-click control, even when uncertain;
- take the correct first-minute actions after clicking, entering credentials, approving MFA, installing software, or sending money; and
- apply enhanced procedures in finance, payroll, HR, executive support, and IT/help-desk roles.

Part I — The threat model employees need

Phishing is a request to cross a trust boundary

Do not define phishing only as a malicious link. The useful definition is **deceptive contact designed to make a person or system take an action that benefits the attacker**. The requested action may be:

- signing in to a fake or adversary-in-the-middle page;
- approving an MFA prompt or disclosing a one-time code;
- scanning a QR code;
- opening an attachment or enabling content;
- installing “support,” “meeting,” or remote-access software;
- sharing a document, customer list, tax record, employee record, source code, API key, or recovery code;
- changing a supplier’s bank account;
- changing an employee’s direct deposit;
- buying gift cards or moving money;
- moving the conversation to text, WhatsApp, a personal email, or an external Teams/Slack tenant; or
- simply replying, which confirms a live target and begins a longer manipulation.

The most durable warning sign is therefore not bad grammar. It is an **unexpected request plus a consequential action**, especially under urgency, secrecy, authority, fear, scarcity, or helpfulness pressure.

What attackers can convincingly fake

Assume an attacker can fake or abuse:

- a display name and profile photo;
- a logo, signature block, legal footer, branded login page, or document-sharing notice;
- polished, idiomatic writing in the recipient’s language;
- caller ID and an SMS sender label;
- a voice sample or synthetic video;
- a recently registered lookalike domain;
- a QR code whose destination is invisible until scanned;
- an external guest identity in a collaboration platform;
- a legitimate cloud-hosting, form, file-sharing, or URL-redirection service;
- a real supplier, customer, colleague, or executive account after compromise;
- information from social media, press releases, calendars, job listings, breached data, and prior email threads; and
- the emotional context of a real event such as a merger, benefits enrollment, tax season, travel, an outage, a new executive, or an urgent deal.

In Microsoft’s Q1 2026 telemetry, 78% of detected email threats were link-based. QR-phishing volume rose from 7.6 million in January to 18.7 million in March, a 146% increase; PDF attachments delivered 65–70% of those QR attacks during the quarter (Microsoft Threat Intelligence, April 2026). Those are Microsoft detections, not an estimate of all global phishing.

What attackers cannot safely replace

Attackers have a harder time defeating a well-designed business process:

- a known phone number retrieved independently from the company directory;
- a callback to a supplier contact already held in the vendor master;
- an in-person check when practical;
- a second authorized person reviewing a payment change;
- a rule that no executive can waive verification by email, chat, voice, or video;
- a FIDO/WebAuthn authenticator that binds authentication to the real site;
- least privilege and transaction limits;
- alerts for new payees, mailbox forwarding rules, risky sign-ins, or new MFA enrollment; and
- rapid employee reporting connected to a rehearsed response team.

This is why verification is a process, not a feeling.

Part II — The 2026 attack landscape

1. AI-written and AI-personalized phishing

Old advice often said to look for spelling errors and unnatural phrasing. Those remain clues when present, but their absence means nothing. Generative systems can rapidly draft clean messages, translate them, imitate tone, summarize public information about a target, and create multiple variants. A polished message can still be malicious; a clumsy message can still come from a compromised legitimate account.

The FBI's 2025 report says AI can create high-quality synthetic content at scale and that it is becoming harder to detect. It specifically identifies official-sounding CEO messages, phishing links, payment directions, and voice cloning as fraud uses (FBI IC3 2025 Annual Report, p. 38). The defensible lesson is not "AI phishing always works" or "most phishing is AI." Available complaint data cannot establish either claim. The lesson is to stop using writing quality as an authentication factor.

Employee rule: Judge the requested action and verify the identity and authority behind it. Do not try to diagnose whether AI wrote the message.

2. Deepfake voice and video vishing

A voice call may imitate an executive, coworker, supplier, family member, or help-desk technician. A video call may combine a synthetic face or voice with stolen images, a compromised account, a low-resolution connection, a prerecorded segment, or a human operator. The absence of visual glitches is not proof of authenticity. Conversely, ordinary network artifacts are not proof of a deepfake.

The FBI has warned that malicious actors use AI-generated voice messages to impersonate senior U.S. officials and build trust before seeking access to accounts; its advice is to verify identity by independently finding a number and, where appropriate, agreeing on a secret word or phrase (FBI public service announcement, May 2025). The general control transfers directly to company vishing: terminate or pause the incoming interaction, then initiate a new contact using a trusted directory entry.

Employee rule: A familiar voice or face may increase urgency, but it does not reduce the need for verification. Never approve money, credentials, MFA, sensitive disclosure, or remote access solely because a caller looks or sounds right.

Safe script: "I'm happy to help. Our policy requires me to call back through the directory and complete verification. I cannot bypass that, including for executives."

3. QR-code phishing ("quishing")

A QR code is a link encoded as an image. It can lead to credential theft, an adversary-in-the-middle sign-in, malware, a fraudulent payment page, or a phone number. It moves inspection from a managed computer to a phone—sometimes a personal device—with a smaller screen and fewer enterprise protections. A code inside a PDF is still an external navigation request.

The FTC advises people not to scan unexpected QR codes in email or text, to inspect the previewed URL, and to contact the supposed sender through a known website or number (FTC Consumer Advice). Microsoft's 2026 data show why this is a current enterprise priority, but volume is less important than the decision rule.

Employee rule: Do not scan an unexpected QR code to sign in, view a document, fix an account, pay an invoice, enroll in benefits, or reset MFA. Navigate through the known application or company portal instead. For a physical code, look for a sticker placed over the original and use the official app or typed address when available.

4. MFA fatigue, push bombing, and token theft

In an MFA-fatigue attack, a criminal who has a password or is attempting a login generates repeated approval prompts, calls the user pretending to be support, or both. The attacker hopes the user will approve to stop the noise. A prompt is not a nuisance to clear; it is a security decision.

CISA describes push bombing as repeated notifications intended to induce acceptance and recommends migrating to phishing-resistant MFA. It identifies FIDO/WebAuthn as the widely available phishing-resistant approach; number matching is a useful interim measure, while SMS and one-time codes remain phishable (CISA, "More than a Password"; CISA phishing-resistant MFA fact sheet). Adversary-in-the-middle pages can also proxy a legitimate sign-in and steal a session token, bypassing non-phishing-resistant MFA; Microsoft documented such a campaign in 2026 (Microsoft Threat Intelligence, May 2026).

Employee rule:

- Approve an MFA request only when you initiated the sign-in, at that moment, on the expected service and device.
- If the prompt is unexpected, select **Deny** or **Report fraud** when offered.
- Do not share a one-time code, number-match value, recovery code, or password with anyone, including "IT."
- Report the unexpected prompt immediately; a password may already be compromised.
- If you approved one, report that fact immediately. Do not wait to see whether anything happens.

5. SMS, messaging-app, Teams, and Slack lures

Text and collaboration messages feel conversational and urgent. Small screens may hide full domains. External-user labels may be overlooked. A compromised colleague's account may carry real history, and an attacker may first build rapport before asking for a device code, file download, screen share, or remote-control session.

Microsoft reports that threat actors abuse Teams messages, calls, meetings, and screen sharing. Observed techniques include external chats, compromised accounts, fake meeting invitations, device-code phishing, malicious ZIP files, and social engineering of help desks and MFA resets (Microsoft Threat Intelligence, October 2025). The same risk model applies to Slack and other collaboration systems even though platform-specific controls differ.

Employee rule: Treat a chat as a message, not as proof of identity. Do not paste a device code supplied by someone else into a login page. Do not install software or grant screen control because an unsolicited "support" contact asks. Verify through the company help portal or directory.

6. Business email compromise

BEC is process fraud conducted through email and, increasingly, phone numbers and virtual-meeting applications. The FBI defines it as a scam targeting organizations or people who work with suppliers or routinely transfer funds, using compromised communications, social engineering, or intrusion to cause unauthorized transfers (FBI IC3 2025 Annual Report, p. 58). BEC may contain no malicious link or attachment. The "payload" is the employee's authorized action.

APWG's Q1 2026 contributing data found that gift cards represented 48% of observed BEC cash-out methods, wire transfers 19%, and payroll diversion 11%. Seventy-two percent of observed attacks used free webmail, while 28% used other domains (APWG Q1 2026 report, pp. 11–12). These are Fortra-observed attacks within the APWG report, not universal market shares.

CEO and executive gift-card fraud

The attacker impersonates a senior leader: "I'm in a meeting," "keep this confidential," "buy ten cards for a client event," "send photos of the codes." The small initial amount may escalate. Secrecy and authority are the mechanism.

Control: Company policy should prohibit gift-card purchases or code transmission requested only by email, text, or chat. Verify through a known number and follow the documented purchasing workflow. Executives must publicly endorse the rule and never criticize an employee for applying it.

Invoice and supplier-payment fraud

The attacker compromises or imitates a supplier, observes a real billing cycle, then changes bank details. The message may arrive in an authentic thread from a real account and include a correct invoice. Sender inspection alone cannot defeat this.

Control: Treat every new payee, bank-detail change, unusual payment route, or urgency request as a high-risk change. Call a preexisting supplier contact using the number already in the vendor master—not the email, invoice, signature, or caller-provided number. Require a second authorized reviewer. Hold and alert on high-risk changes.

Payroll diversion

The attacker impersonates an employee and asks HR/payroll to change direct deposit, often using personal email and an explanation about account access.

Control: Accept bank changes only through the authenticated HR system or a documented identity-proofing process. Notify the employee through a previously registered channel and apply a cooling-off or enhanced review period where lawful and practical. A manager's emailed approval is not identity proof.

Executive and legal-request fraud

The attacker claims secrecy is required for an acquisition, litigation, tax payment, regulatory issue, or confidential vendor. They may impersonate outside counsel or place a deepfake call.

Control: Confidentiality never cancels verification or dual authorization. Escalate through the legal and finance paths already on file. "The CEO said to bypass policy" is itself a red flag.

Part III — How to inspect a message

Inspection is a triage tool, not a guarantee. A message with no visible red flags can come from a compromised account. Use inspection to decide whether to stop and verify; always verify high-risk requests regardless of appearance.

1. Inspect the actual sender

Separate four concepts:

- **Display name:** the friendly label, freely chosen and easy to spoof.
- **From address:** the visible mailbox address.
- **Reply-To:** where replies may actually go; it can differ from From.

- **Sending infrastructure/authentication:** technical evidence such as SPF, DKIM, and DMARC, normally evaluated by mail systems and security teams. A message can pass authentication yet still be malicious if it comes from a compromised account or an attacker-controlled domain.

On desktop, expand the sender details. On mobile, tap the sender name or details. Compare the entire address with the expected company or supplier domain. Do not rely on the avatar or display name.

Worked sender example

```
Display name: Maya Chen, Chief Financial Officer
From: maya.chen@northwind-finance.example
Reply-To: urgent.board.request@consultant-mail.example
Subject: Confidential acquisition payment
```

Assume the company's real documentation domain is `northwind.example`. This message uses `northwind-finance.example`, a separate domain, and diverts replies again. Both are reserved illustrative domains; they do not represent real organizations.

Red-flag callouts

1. The display name proves nothing.
2. The registrable domain is not `northwind.example`.
3. Reply-To differs from From.
4. "Confidential acquisition" creates secrecy and authority pressure.
5. A payment request crosses a financial trust boundary and requires independent verification even if every address were correct.

2. Find the registrable domain

Read domain names from right to left. The important ownership boundary is usually the registrable domain immediately before the public suffix.

```
login.northwind.example
```

Here, `northwind.example` is the base used in this fictional example; `login` is a subdomain.

```
northwind.example.secure-review.invalid
```

Here, `secure-review.invalid` is the controlling domain. The words `northwind.example` are only labels placed to its left.

```
northwínd.example
```

The `í` is not `i`. Internationalized-domain homographs may render deceptively. Browsers and mail clients may show an ASCII `xn--…` form in some contexts. Do not manually "correct" a suspicious spelling and proceed; navigate from a bookmark or company portal.

```
northwind-payments.example
northwind.example
```

These are different domains. Hyphens, added words, omitted letters, doubled letters, swapped letters, and different endings create distinct registrations.

Important limitation: A perfect match still does not prove the request is safe. The account, supplier, or tenant may be compromised.

3. Inspect a link without opening it

On a desktop, hover over the link and read the destination shown by the client. On a touch device, a long press may show a preview, but behavior varies and can accidentally open a link; when uncertain, do not interact and report. Expand shortened links only with an approved company tool, not a public service that may leak sensitive URLs.

Worked link example

Message text:

```
Review the secure document:
https://portal.northwind.example/document/4281
```

Hover destination:

```
https://northwind.example.review-center.invalid/session/4281
```

The visible text is a label; the destination controls where the click goes. The actual controlling domain is `review-center.invalid`, not `northwind.example`.

HTTPS and the padlock

`https://` means the connection to that site is encrypted and the certificate is valid for the displayed host. It does **not** mean the site is honest, affiliated with the brand it depicts, or safe to enter credentials into. Criminal sites routinely use HTTPS.

Redirects and trusted platforms

A link may begin on a real cloud, form, advertising, or redirection service and then send the user elsewhere. A legitimate first domain does not bless the final destination. If the message says a document or account needs attention, open the known app from a bookmark, launcher, or typed address and find the item there.

4. Inspect context and requested action

Ask:

- Did I expect this contact?
- Is the request normal for this person and my role?
- Does it create urgency, secrecy, fear, authority, curiosity, or reward?
- Am I being asked to sign in from a link rather than through the normal app?
- Does it ask for a password, MFA code, recovery code, API key, or personal data?
- Does it change money, bank details, direct deposit, tax data, or a delivery address?
- Does it move me to a personal device or new channel?
- Does it ask me to install software, run a command, share my screen, or grant control?
- Can I verify using contact information I already trust?

A single cue rarely proves phishing. A consequential request is enough to require verification even without cues.

Part IV — Realistic walkthroughs for live training

Walkthrough A: “Flawless” benefits email

```
From: People Operations <benefits@northwind-people.example>
Subject: Action required by 4:00 p.m.: dependent coverage reconciliation

Hello Jordan,

As part of our midyear benefits reconciliation, please confirm the
dependents attached to your medical plan. This review is required to
prevent interruption of coverage.

Review dependent coverage

The form will close at 4:00 p.m. today. If you have already completed
the review, no further action is required.

People Operations
```

Hover destination:

```
https://sso.northwind.example.hr-review.invalid/continue
```

Callouts for the presenter

- The grammar and tone are professional. That is not evidence of legitimacy.
- Benefits and fear of losing coverage are contextually relevant; NIST research shows that alignment with work context affects detection difficulty (NIST workplace study).
- The sending domain is a plausible lookalike, not the illustrative company domain.
- The link’s controlling domain is `hr-review.invalid`.
- The deadline suppresses deliberation.
- Safe action: open the HR portal from the company launcher or bookmark, check for the task, and report the message.

Walkthrough B: Compromised supplier, real invoice thread

```
From: accounts@blue-harbor-supplies.example
Subject: RE: Invoice 78421 — June services

Hi Priya,

Following our bank migration, please use the attached remittance
instructions for invoice 78421 and all future payments. We need this
released today to avoid a service interruption.

Regards,
Sam
```

Assume the sender, invoice number, signature, and thread history are all real because the supplier mailbox was compromised.

Callouts

- There may be no sender or grammar anomaly.
- The high-risk event is the bank-detail change.
- An attachment could be genuine-looking and still carry fraudulent instructions.

- Replying in the same thread reaches the attacker.
- Calling the number in the new instructions or signature may reach the attacker.
- Safe action: stop the payment; call the previously registered supplier contact in the vendor master; require dual review; report the email and possible supplier compromise.

Walkthrough C: CEO gift cards by text

```
Unknown number:
Hi, it’s Elena. I’m heading into the board meeting and need a quick
favor. Can you buy eight $200 gift cards for client recognition?
Keep this between us until the announcement. Send photos of the codes.
I’ll reimburse the expense today.
```

Callouts

- Unknown or spoofed number.
- Authority plus helpfulness.
- Secrecy blocks consultation.
- Urgency bypasses purchasing controls.
- Gift-card codes are cash-like and irreversible.
- Safe action: do not reply or purchase; contact Elena through the directory or executive assistant; report the text. Policy applies even if the voice later sounds like Elena.

Walkthrough D: Teams “IT support”

```
External — Help Desk Support

We detected sync failures on your laptop. Join this call now so we can
repair your profile before files are lost. Download the support client
from the link below and share the six-digit device code in the call.
```

Callouts

- External identity presenting as internal support.
- Fear of data loss and immediate deadline.
- Unsolicited software download.
- Device-code request: an attacker may be trying to authenticate their device using the employee’s authorization.
- Safe action: end the chat/call; open the approved help portal independently; contact the help desk using its published number; report the conversation.

Walkthrough E: QR-code document

```
Subject: Secure payroll document available

This encrypted document cannot be displayed on a desktop browser.
Scan with your mobile device to authenticate and view.

[QR CODE]
```

Phone preview:

```
https://northwind-payroll.example.login-gateway.invalid
```

Callouts

- The instruction deliberately moves the user off the managed desktop.
- “Cannot be displayed” tries to make the unusual workflow sound necessary.
- The preview contains a brand-like label, but the controlling domain is `login-gateway.invalid`.
- Safe action: do not scan; use the normal payroll portal; report the message.

Walkthrough F: MFA fatigue plus voice call

Sequence:

1. An employee receives five unexpected “Approve sign-in?” prompts.
2. A caller claiming to be IT says: “We are testing your account. Accept the next prompt so I can stop the alerts.”
3. The caller knows the employee’s name, title, and manager.

Callouts

- The unexpected prompts may mean the password is already known.

- IT should not ask a user to approve an uninitiated login or disclose a code.
- Public or breached data can explain the caller's knowledge.
- Safe action: deny/report the prompt, hang up, call the official help desk, and report immediately. If one was approved, say so explicitly so responders can revoke sessions and investigate.

Walkthrough G: Payroll diversion

```
From: Jordan Lee <jordan.lee.personal@personal-mail.example>
To: Payroll
Subject: Direct deposit update before today's cutoff

I'm locked out of the HR portal while traveling. Please replace my
current account with the attached form. My manager is copied and has
approved. I need this done before payroll closes.
```

Callouts

- Personal email and claimed account lockout.
- Payroll cutoff creates urgency.
- Copying or imitating a manager does not prove employee identity.
- The attachment may contain a mule account.
- Safe action: use the authenticated HR change process or established identity-proofing exception; notify Jordan through a previously registered channel; never use contact data supplied in the request.

Part V — Stop. Think. Verify. Report.

STOP

Pause before clicking, scanning, opening, installing, signing in, approving, sharing, or paying. A pause of seconds is often enough to switch from reflex to procedure. Do not let a sender's stated deadline override company controls.

THINK

Name the requested action and its consequence:

- "This asks me to enter my company password."
- "This asks me to approve a login I did not start."
- "This changes a bank account."
- "This asks for confidential employee data."
- "This asks me to install software and grant remote control."

Then identify the pressure: urgency, secrecy, fear, authority, scarcity, curiosity, reward, or helpfulness.

VERIFY

Verification must be **out of band**: use a separate, trusted route that the suspicious contact did not supply or control.

Good verification:

- call the directory number you already use;
- open the known application from a bookmark or company launcher;
- use the supplier contact and number already stored in the vendor master;
- ask in person;
- initiate a new message to the known account rather than replying to the suspicious thread;
- contact the help desk through the service portal;
- obtain the required second approval from an independently authenticated approver.

Bad verification:

- reply "Is this really you?" to the same message;
- call the number in the suspicious email, invoice, attachment, text, or caller ID;
- use a link or QR code in the message;
- ask the caller to prove identity with facts available on social media;
- accept a video face or familiar voice as proof;
- let the requester choose the verifier;
- treat a copied manager as approval; or
- allow "confidential" or "the CEO authorized it" to bypass control.

> ***Universal money-and-credential rule***
>
> *Any unexpected request involving money, payment instructions, payroll, gift cards, credentials, MFA, recovery codes, sensitive data, software installation, or remote access must be verified through a trusted channel chosen independently of the request. No person, including the CEO, can waive this rule inside the requesting communication.*

REPORT

Report when suspicious, when uncertain, and after a mistake. Do not merely delete. Reporting gives defenders the sender, headers, URLs, attachment metadata, recipients, and timing needed to protect others.

Company reporting procedure — configure before delivery

1. In Outlook/Gmail, select the message and click **[Report Phishing button name]**.
2. In Teams/Slack, use **[Message menu → Report]** and select **[Security/Phishing]**. If the platform does not preserve sufficient evidence, use **[security intake channel]**.
3. For SMS, QR codes, calls, voicemail, personal email, or physical mail, send a screenshot and details to **[security portal or monitored address]** or call **[security/help-desk number]**.

4. For an accidental click, credential entry, MFA approval, software installation, sensitive-data disclosure, or payment, mark it **URGENT** and call **[24×7 incident number]** immediately after submitting the report.
5. Preserve the message and wait for guidance. Do not forward a suspicious attachment or link to coworkers. Do not investigate it yourself.

If a one-click report button is not yet deployed, the company has not completed this program's minimum operating requirements. CISA's guidance for security-awareness programs emphasizes making reporting easy and giving users feedback; NIST's current awareness-program guidance likewise treats behavior, culture, and usable reporting mechanisms as program concerns, not just course completion (NIST SP 800-50 Rev. 1, 2024).

Part VI — No-blame reporting culture

The policy

The organization will:

- thank employees for prompt reports, including false positives;
- accept "I'm not sure" as a valid reason to report;
- separate an honest mistake and prompt report from reckless, malicious, or repeated policy evasion;
- never publish individual simulation failures, shame teams, or use a leaderboard of "clickers";
- give reporters timely closure when feasible;
- make the reporting path easier than deleting or ignoring;
- train managers to respond first with "Thank you for telling us quickly";
- design simulations around realistic work risks without exploiting personal hardship, compensation, layoffs, illness, or other unnecessarily harmful themes;
- collect only the data needed for improvement and response; and
- measure reporting and response quality, not just clicks.

Why this is operationally necessary

Shame delays disclosure. In a live incident, minutes matter for revoking sessions, resetting credentials, isolating a device, recalling messages, blocking infrastructure, warning peers, and attempting fund recovery. The FBI says rapid complaint filing supports its Financial Fraud Kill Chain process (FBI IC3 2025 Annual Report, p. 9). A punitive culture may produce a cosmetically low click rate while making real incidents less visible.

Research on employee motivation in phishing interventions finds that program communications, feedback, beliefs about usefulness, and social context influence participation and reporting intentions; it argues for designing interventions around employee value rather than treating users as a problem (Chen et al., SOUPS 2024). This does not prove that a specific gratitude script reduces breach losses, but it supports a human-centered approach and aligns with the operational need for early warning.

Manager response script

> *"Thank you for reporting it quickly. You did the right thing. Tell me what happened and what actions you took; you are not in trouble for giving us the facts. Security will help contain it."*

Avoid:

- "How could you fall for that?"
- "Everyone knows this."
- "Don't tell anyone."
- "Just change your password and forget it."
- "We'll see whether it becomes a problem."

Part VII — The first minutes after a bad click

Speed and accuracy are more valuable than embarrassment or self-investigation.

If you clicked but did not enter information

1. **Stop interacting.** Close the page or application. Do not download anything, accept notifications, grant permissions, or call numbers shown on the page.
2. **Report immediately** with the one-click button or urgent channel. State that you clicked and give the approximate time.
3. **Follow security's instructions.** Do not erase history, delete files, power off, or factory-reset unless directed; responders may need evidence. If the device is clearly executing something suspicious or security directs isolation, disconnect from networks using the

approved procedure.

4. **Do not assume “nothing happened.”** A page can redirect, fingerprint a device, trigger a download, or begin a session even without visible data entry. Conversely, do not claim that every click causes compromise; responders will assess it.

If you entered a password

1. **Report and call the incident line immediately**, from a safe device if instructed.
2. **Go to the real account through a bookmark or company portal** and change the exposed password, following security guidance.
3. **If the password was reused anywhere, disclose that privately to responders and change it on every reused account**, starting with email and financial accounts. Use unique passwords stored in the approved password manager.
4. **Let IT revoke sessions/tokens and inspect mailbox rules, MFA methods, devices, and sign-ins.** A password change alone may not invalidate an attacker’s session.
5. **Do not approve subsequent MFA prompts.**

If you entered an MFA code, approved a prompt, registered a passkey/security key, or pasted a device code

1. **Deny/report any further prompts.**
2. **Call security immediately and say exactly which action occurred.**
3. **Expect session revocation, credential reset, and review of authentication methods and sign-in logs.**
4. **Do not remove evidence or attempt to “undo” enrollment without direction**, because responders need to determine what changed.

If you opened an attachment, ran a file, command, macro, installer, or remote-support tool

1. **Stop interacting and contact security by phone.**
2. **Disconnect the device from Wi-Fi/Ethernet only if the organization’s incident card directs this or security instructs it.** Do not disconnect a clinical, industrial, safety-critical, or life-safety system without its specific procedure.
3. **Leave the device powered on unless responders instruct otherwise.**
4. **Tell responders the filename, time, prompts accepted, commands run, and whether remote control was granted.**
5. **Use another known-safe device for sensitive communications if directed.**

If you disclosed sensitive data

Report exactly what categories and whose data were involved: customer, health, employee, tax, government ID, payment card, bank, source code, secret, or regulated data. Do not send another copy of the data in the report unless the approved secure channel requests it. Security, privacy, legal, and compliance teams will determine notification and containment duties.

If money, gift-card codes, or payment details were sent

1. **Call the organization’s finance escalation and security lines immediately.**
2. **Contact the bank or payment provider through established channels to request a hold, recall, or freeze.**
3. **Preserve the transaction details, messages, invoices, account numbers, timestamps, and verification steps.**
4. **For U.S.-connected incidents, report promptly to IC3 at the real address reached independently: [ic3.gov](https://www.ic3.gov/).** Law enforcement and counsel should guide additional reporting.
5. **Do not contact the attacker to negotiate or warn them.**

The precise isolation and recovery steps are organization-specific. The printed quick card must contain the correct phone numbers and must be tested quarterly.

Part VIII — Do / don’t comparisons

Situation	Do	Don’t
Unexpected sign-in notice	Open the known service from a bookmark and inspect account activity	Use the email’s “secure account” link
MFA prompt you did not initiate	Deny/report it and contact security	Approve to make prompts stop
Executive asks for gift cards	Call back using the directory and follow purchasing policy	Trust the display name, voice, or urgency
Supplier changes bank details	Call the preexisting vendor-master contact; use dual approval	Reply in the thread or call the new invoice number
Employee changes direct deposit	Require authenticated HR workflow or approved proofing	Accept a form from personal email
QR code in a document	Navigate to the known portal independently	Scan because the PDF looks branded
Teams/Slack “IT” contact	Use the official help portal/number	Install tools, paste device codes, or grant control
Suspicious message, uncertain	Report it	Delete silently or poll coworkers by forwarding it
Clicked a bad link	Stop, report immediately, tell the full story	Hide it, wait, or investigate the site
Caller sounds like a leader	End the incoming interaction and initiate a directory callback	Use voice or video as the sole identity check

Part IX — Role-based modules

Finance, accounts payable, treasury, and procurement

Mandatory controls

- Independent callback for every new payee and every bank-detail change.
- Use only contact data already in the vendor master or independently sourced from a trusted contract—not the change request.
- Segregation of duties: requester, verifier, approver, and releaser should be separated according to risk.
- Dual authorization for high-risk or out-of-pattern payments.
- Transaction limits, new-payee holds, and alerts.
- Verification log recording who was contacted, through which trusted channel, when, and what was confirmed.
- Escalation for urgency, secrecy, unusual geography/currency, split payments, overpayments, refunds to a new account, and last-minute changes.
- Periodic vendor-master review and change notifications to established supplier contacts.

Scenario drill: A real supplier account sends correct invoice history but requests new bank details. The expected outcome is “hold, independently verify, dual approve, report”—not “spot a typo.”

Payroll and HR

- Direct-deposit and tax-detail changes through an authenticated portal.
- Strong identity proofing for exceptions; no knowledge questions based on public data alone.
- Confirmation through a previously registered channel.
- Alerts to employees when bank, address, phone, MFA, or tax settings change.
- Separation between personnel-data access and payroll-release authority.
- Extra care with new hires, executives, remote workers, and pre-payroll cutoff urgency.
- Never email full tax forms, national identifiers, or bank data in response to an unexpected request.

Executives and executive assistants

- Publicly endorse callback and dual-control rules.
- Expect employees to decline unverified requests.
- Do not create exceptions by saying “I’ll never ask this, except when urgent.”
- Minimize public exposure of travel, reporting lines, deal timing, and personal voice/video where feasible, while recognizing that exposure cannot be eliminated.
- Use phishing-resistant MFA and separate privileged functions.
- Prearrange an escalation route for truly confidential transactions that still preserves independent verification.

IT and help desk

- Never ask users for passwords, one-time codes, recovery codes, or approval of a login the user did not initiate.
- Authenticate callers before resets; resist urgency and seniority pressure.
- Apply stronger proofing for MFA reset, phone-number change, passkey enrollment, device registration, and privileged access.
- Alert on and review new authentication methods, risky sign-ins, impossible travel, inbox rules, OAuth grants, and device-code flows.
- Restrict external collaboration and remote-control tools according to business need.
- Publish exactly what legitimate support contact looks like.

Sales, customer support, and communications

- Expect lures using contracts, RFPs, complaints, shared files, media requests, and customer urgency.
- Open documents in protected workflows.
- Do not bypass authentication to “help” a customer.
- Verify requests to change customer account owners, delivery addresses, refund destinations, or administrator access.

Developers and administrators

- Treat requests for secrets, tokens, package installation, repository access, OAuth consent, CI/CD changes, and cloud invitations as credential or access requests.
- Never paste commands from an unsolicited message into a shell.
- Use hardware-backed, phishing-resistant authentication and separate administrative accounts.
- Verify package names and repository invitations independently.

Part X — Delivery plan

Core session: 60–75 minutes

Time	Segment	Presenter action
0–5 min	Purpose and no-blame promise	Read the opening script; show reporting button and urgent number
5–12 min	Modern threat landscape	Explain cross-channel attacks; use current FBI, APWG, Verizon, and Microsoft evidence
12–20 min	The trust-boundary model	Ask participants to name consequential actions, not spelling errors
20–35 min	Sender, domain, and link inspection	Work through examples A and link parsing live
35–47 min	BEC, deepfake, QR, MFA, and chat	Work through examples B–G; rehearse callback language
47–55 min	Stop–Think–Verify–Report	Have participants say the four steps and demonstrate reporting
55–63 min	After-click response	Present the branching first-minute playbook
63–70 min	Knowledge check	Run the quiz; discuss answers
70–75 min	Commitment and contacts	Show the final quick card; invite uncertain reports

Facilitation notes

- Replace every bracketed placeholder before delivery.
- Demonstrate the real one-click report control in the production mail client.
- Display the official help-desk and incident numbers long enough for participants to save them.
- Tell participants all example domains are reserved and fictional.
- Do not send live links in training materials that imitate a login page.
- Encourage participants to challenge a simulated executive request aloud.
- Provide accessible text descriptions for every image and captioned alternatives for audio/video examples.
- Do not use an actual employee’s failure as a case study without informed consent and careful anonymization.

Standalone slide-deck blueprint

The later slide deck can stand alone by mapping one primary idea to each slide:

1. Title and no-blame promise
2. Today’s objective: Stop–Think–Verify–Report
3. Phishing is a request to cross a trust boundary
4. What can be faked
5. What a safe process protects
6. Current evidence: FBI/APWG/Verizon, with measurement caveats
7. AI: flawless language is not authentication
8. Deepfake voice/video: callback rule
9. QR codes are links
10. MFA fatigue and AiTM
11. Teams/Slack/SMS lures

12. BEC and the $3.047 billion IC3-reported loss figure
13. Gift-card fraud walkthrough with callouts
14. Invoice-change walkthrough with callouts
15. Payroll-diversion walkthrough with callouts
16. Actual sender versus display name
17. Read domains right to left
18. Hover/preview links; HTTPS is not trust
19. Flawless benefits email walkthrough
20. Compromised legitimate account: why cues are insufficient
21. Stop
22. Think
23. Verify out of band: good versus bad
24. Report: live demonstration
25. No-blame promise
26. First minutes after a click
27. First minutes after credential/MFA exposure
28. First minutes after software execution or payment
29. Technical controls are the safety net

30–35. Knowledge-check questions

36. Answers and rationale
37. Quick-reference card and contacts

Every walkthrough slide should show the full fictional message on the left and numbered callouts on the right. Notes should contain the relevant source link and the safe behavior. Avoid relying on speaker notes for the reporting procedure; it must appear on the slide.

Standalone podcast/audio-lesson blueprint

Target 25–35 minutes:

1. **Cold open:** a flawless CEO voice asks for a payment; pause before revealing it may be synthetic.
2. **Reframe:** phishing is a consequential request, not a spelling test.
3. **Threat tour:** email, SMS, QR, Teams/Slack, voice/video, MFA fatigue, and BEC.
4. **Audio walkthroughs:** narrate the gift-card, invoice, payroll, and support scenarios. Spell out fictional domains slowly and explain “read right to left.”
5. **Habit rehearsal:** repeat “Stop. Think. Verify. Report.” Use a call-and-response pause.
6. **After-click branches:** clicked only; entered password; approved MFA; ran software; sent money.
7. **No-blame close:** report even when unsure; fast honesty is a security control.
8. **Contacts:** record company-specific reporting details at both the midpoint and end.

Because listeners cannot hover links or see sender details, the audio version must teach navigation through known apps and independent callbacks as the default. A transcript should accompany it for accessibility.

Part XI — Continuous program: beyond the annual course

Recommended cadence

The annual course remains useful for onboarding, policy acknowledgement, and a common baseline, but it is not the behavior-change mechanism by itself.

- **Onboarding, day 1–7:** 20–30-minute core lesson, reporting-button practice, and role-specific rules.
- **Monthly:** 3–5-minute micro-lesson tied to current organizational threats; one behavior only.
- **Every 4–8 weeks:** a respectful, risk-based simulation or scenario rehearsal, with cadence adjusted to fatigue and operational risk.
- **Quarterly:** role-based practice for finance, payroll/HR, help desk, executives, and privileged users.
- **After a material campaign:** a rapid advisory using sanitized real examples and a single action.

- **Annually:** full program refresh, control demonstration, and tabletop exercise involving security, communications, finance, HR, legal, and leadership.

This cadence is a program recommendation, not a universally proven optimum. The literature does not establish one frequency that works for every organization. Continuous, short practice is favored because threats and business context change, reporting needs rehearsal, and annual-course recency alone showed no protective relationship in the 2025 UCSD field experiment. The recommendation should be tested locally with control groups or phased rollouts where ethical and practical.

Design principles for micro-training

- Teach one decision in under five minutes.
- Use the employee's real tools and approved process.
- Vary channels, not only email.
- Calibrate difficulty with NIST's Phish Scale concepts: observable cues plus premise alignment to the recipient's context (NIST Phish Scale User Guide, TN 2276).
- Include non-click scenarios such as replying, calling a number, changing a bank account, approving MFA, or disclosing data.
- Give everyone learning, not only those who click. The UCSD study found that click-triggered embedded training reached only a minority in each campaign and was often barely viewed.
- Keep feedback short, specific, and immediate enough to connect action and lesson, while recognizing that field evidence on the best timing is still developing.
- Reward reporting and verification.
- Avoid tricks unrelated to the intended behavior.
- Never promise that trained employees will detect every attack.

Simulation charter

Before simulations begin, leadership, HR, legal, privacy, labor/employee representatives where applicable, and security should approve a charter:

- **Purpose:** rehearse reporting and verification; evaluate process and controls.
- **Not purpose:** embarrass, rank, or punish.
- **Scope:** realistic work channels and role-based risks.
- **Prohibited themes:** exploitative promises or threats involving layoffs, bonuses, health crises, bereavement, or other sensitive topics unless there is a compelling risk reason and appropriate review.
- **Data:** minimum necessary, retention limit, access control, aggregation for management.
- **Feedback:** respectful, specific, accessible, and available to all.
- **Escalation:** supportive coaching for repeated risky behavior; investigate process barriers before attributing fault.
- **Controls:** simulations must not collect real passwords, execute code, or create unsafe habits.
- **Measurement:** include report rate, time to report, correct verification, and help-desk/security response—not only failure rate.

What the evidence does and does not support

Evidence suggesting benefit

Longitudinal workplace and large-enterprise studies have observed declining click rates across repeated exercises. For example, Gordon et al. observed overall click rates decrease across 20 campaigns, and Siadati et al. analyzed 115,080 simulated messages sent to 19,180 participants to develop less biased evaluation methods (Gordon et al., 2019; Siadati et al., USENIX CSET 2017). Repeated exposure can provide practice, reveal process weaknesses, and generate reporting telemetry.

Evidence questioning strong claims

Declining raw click rates are not necessarily causal evidence of training. Lures differ in difficulty; recipients remember repeated templates; filters change; workforce composition changes; and people who do not click may receive no instruction. NIST's Phish Scale was created because click and report rates require context (Steves, Greene, and Theofanos, *Journal of Cybersecurity*, 2020).

The 2025 UCSD randomized study is especially important because it included control groups and varied content. It found:

- no clear relationship between annual-training recency and simulation failure;
- an average two-percentage-point reduction from embedded training;
- lure failure rates ranging from roughly 1–2% to above 30%, larger than the measured training effect;
- more than 56% of participants clicked at least once across the study;
- more than half of embedded-training visits ended in under ten seconds; and

- fewer than 24% formally completed the training (UCSD/IEEE S&P 2025).

The study was conducted in one healthcare organization over eight months and measured simulation behavior, not breach losses. Its results should constrain broad promises, not end all awareness work. It indicates that merely sending simulations and redirecting clickers to content is a weak intervention.

The 2019 healthcare study's targeted mandatory course also failed to produce a substantial additional reduction among high-risk employees. It was observational rather than randomized and its "offender" categorization selected people based on past clicks, but it reinforces the need to evaluate remedial training rather than assume it works (Gordon et al., 2019).

Honest conclusion

It is too strong to say "continuous micro-training and simulations beat annual courses" in every setting. The defensible conclusion is:

- annual training alone has weak evidence as a behavioral defense;
- short, repeated, contextual practice is a plausible and operationally useful way to maintain awareness, rehearse reporting, and adapt to changing threats;
- simulations can measure and exercise parts of the system, but raw click rates are confounded;
- program benefit depends on engagement, relevance, reporting culture, feedback, and surrounding controls;
- randomized or well-controlled local evaluation is preferable; and
- no awareness intervention can substitute for phishing-resistant authentication and sound business processes.

Part XII — Technical and process controls: the safety net

Employee vigilance is variable by design: people are hired to serve customers, deliver care, build products, and pay legitimate invoices. Security must make the safe path easy and constrain the impact of an inevitable mistake.

Identity

- Deploy FIDO2/WebAuthn passkeys or security keys, prioritizing administrators, executives, finance, HR, help desk, remote access, and email.
- Disable legacy authentication.
- Use number matching as an interim control where phishing-resistant MFA is unavailable.
- Reduce reliance on SMS and phishable OTP.
- Apply conditional access, managed-device requirements, risk-based sign-in controls, and sensible session lifetimes.
- Protect MFA enrollment and recovery as privileged events.
- Alert and investigate unexpected prompts, new methods, device-code authentication, risky OAuth consent, and session anomalies.
- Give users a clear way to report fraud from the prompt.

CISA calls phishing-resistant MFA the gold standard and specifically recommends FIDO/WebAuthn because origin binding blocks authentication to an impostor site (CISA fact sheet; CISA "More than a Password"). "Phishing-resistant" is not "fraud-proof": session theft, endpoint compromise, recovery-process abuse, malicious OAuth consent, and authorized payment manipulation still require controls.

Email and domains

- Enforce SPF, DKIM, and DMARC with monitoring and a staged move toward rejection.
- Use anti-impersonation, lookalike-domain, URL, attachment, QR, and BEC detection.
- Label external mail clearly, while avoiding habituation to noisy banners.
- Block automatic external forwarding and alert on suspicious mailbox rules.
- Protect high-value display names and domains.
- Monitor newly registered lookalikes and abuse of company brands.
- Provide a one-click reporter that submits the original message and metadata.
- Automate post-report search, purge, detonation, and blocking where safe.

Email authentication reduces domain spoofing; it does not validate message intent and cannot prevent abuse of a compromised authorized account.

Collaboration, voice, and mobile

- Restrict external federation, guest chat, unmanaged applications, file transfer, and remote control to business need.
- Make external identities conspicuous.
- Log and alert on anomalous chats, device-code flows, new app consent, and external file delivery.

- Publish support-contact norms.
- Protect telecom accounts against SIM swaps and port-out fraud.
- Provide a reporting path for texts, calls, voicemail, and QR codes.

Endpoints and browsers

- Managed browsers, safe-link analysis, DNS/web filtering, EDR, application control, macro restrictions, sandboxing, least privilege, and rapid isolation.
- Prevent ordinary users from installing unapproved remote-access tools.
- Keep operating systems, browsers, mobile devices, and document readers updated.
- Use protected viewing and content disarm/reconstruction where risk warrants.

Finance, payroll, and data

- Independent verification, segregation of duties, transaction limits, and dual approval.
- New-payee and bank-change holds.
- Behavioral analytics and alerts for unusual amounts, destinations, currencies, timing, or approvers.
- Authenticated self-service for payroll changes.
- Data-loss prevention and least privilege for sensitive exports.
- A rehearsed bank and law-enforcement escalation plan.

Response

- A monitored 24×7 urgent channel appropriate to organizational risk.
- Automated acknowledgement to reporters.
- Playbooks for clicked link, credential theft, token theft, malicious OAuth, endpoint execution, BEC, payroll, and data disclosure.
- Session revocation—not password reset alone—where identity compromise is suspected.
- Mailbox rule, delegated access, app consent, authentication method, and sign-in review.
- Rapid search and purge across recipients.
- Finance/bank recall and IC3 escalation paths.
- Tabletop exercises that include executives and third parties.

Part XIII — Measurement and governance

Measure outcomes, not theater

Use a balanced scorecard:

Leading indicators

- percentage of users who can access the report button;
- median and 90th-percentile time to report;
- report rate for simulations, adjusted for difficulty;
- correct out-of-band verification rate in role-based exercises;
- percentage of high-risk payments with documented independent verification;
- phishing-resistant MFA coverage by role;
- time to acknowledge a report;
- time to search/purge matching messages;
- time to revoke sessions after a credential/MFA report;
- false-positive reports handled respectfully;
- training engagement and demonstrated knowledge, not completion alone.

Lagging indicators

- confirmed account takeover;
- unauthorized payments and recovered funds;
- payroll diversion;
- malware execution;
- sensitive-data disclosure;
- dwell time from user action to containment;
- incidents first detected by employee reporting;
- repeated process-control bypass.

Interpret simulation metrics correctly

Do not compare teams using raw click rate unless lure, timing, exposure, role, and context are comparable. Score lure difficulty with the NIST Phish Scale, preserve a control or phased cohort where practical, and report confidence intervals and sample sizes. Separate:

- delivered;
- opened, if measurement is lawful and reliable;
- clicked/scanned;
- entered simulated data;
- reported;
- time to report;
- verified independently; and
- security-team handling.

A rising report rate can increase the security queue and may include false positives; that can be a sign of trust, not failure. Tune automation and staffing instead of discouraging reports.

Twelve-month implementation roadmap

Days 0–30

- Appoint executive sponsor and program owner.
- Publish no-blame policy.
- Deploy or validate one-click reporting.
- Establish urgent incident number and after-click card.
- Inventory payment, payroll, MFA-recovery, and vendor-change workflows.
- Baseline phishing-resistant MFA coverage.

- Run the core session for security, HR, legal, communications, finance, and leadership first.

Days 31–90

- Deliver core training to all personnel.
- Implement finance/payroll callback and dual-control standards.
- Begin role-based sessions.
- Instrument report acknowledgement, time to triage, and response.
- Rate existing simulation templates for difficulty and harmful themes.
- Prioritize FIDO/WebAuthn for high-risk roles.

Months 4–6

- Start monthly micro-lessons and respectful multichannel scenarios.
- Exercise MFA fatigue, QR, Teams/Slack, invoice, payroll, and deepfake callback cases.
- Test message search/purge, session revocation, and bank escalation.
- Review false positives and employee feedback.

Months 7–12

- Expand phishing-resistant MFA.
- Conduct a cross-functional BEC/deepfake tabletop.
- Evaluate cohorts and difficulty-adjusted outcomes.
- Remove ineffective content and friction in reporting.
- Audit vendor-master and payroll controls.
- Report program outcomes and limitations to leadership.

Part XIV — Knowledge check

Questions

1. A message from the correct supplier address asks you to pay a real invoice to a new bank account. What is the safest next step?

- A. Reply and ask whether the change is real.
- B. Call the number in the new invoice.
- C. Hold payment and call the preexisting vendor-master contact; follow dual approval.
- D. Pay because the sender and invoice are genuine.

2. You receive an MFA approval prompt while reading email, but you are not signing in. What should you do?

- A. Approve it once.
- B. Deny/report it and contact security immediately.
- C. Ignore repeated prompts indefinitely.
- D. Tell the caller the number shown.

3. Which domain controls this fictional URL: `https://payroll.northwind.example.secure-check.invalid/update`?

- A. `northwind.example`
- B. `payroll.northwind.example`
- C. `secure-check.invalid`
- D. `update`

4. A video caller looks and sounds exactly like the CEO and requests a confidential wire. What should you do?

- A. Proceed because video is strong proof.
- B. Ask for a personal fact.
- C. Independently call through the directory and follow payment controls.
- D. Ask the caller to send an email.

5. You clicked a suspicious link but entered nothing. What should you do first?

- A. Wait for symptoms.

- B. Investigate the website.
- C. Stop interacting and report the click immediately.
- D. Delete the email.

6. An unexpected QR code says it opens a protected HR document. What is safest?

- A. Scan it on a personal phone.
- B. Open the known HR portal independently and report the message.
- C. Scan it because QR codes cannot execute links.
- D. Forward it to a colleague to test.

7. When is it appropriate to report a suspicious message?

- A. Only when certain it is malicious.
- B. Only after clicking.
- C. When suspicious or unsure, and immediately after any interaction.
- D. Only if it asks for money.

8. Why is flawless grammar not a safety signal?

- A. Legitimate senders never use good grammar.
- B. Generative tools and careful attackers can create polished messages.
- C. Grammar has no relationship to communication.
- D. Mail filters correct attackers' writing.

9. What makes verification "out of band"?

- A. It uses a different trusted route selected independently of the request.
- B. It asks the sender twice.
- C. It uses the phone number in the email signature.
- D. It copies a manager.

10. Which statement is most accurate?

- A. Training makes technical controls unnecessary.
- B. A low simulation click rate proves the organization is secure.
- C. Training is rehearsal; phishing-resistant MFA, business controls, and response limit inevitable mistakes.
- D. Employees are the firewall.

Answers and rationale

1. **C.** A real account can be compromised. The process event—new bank details—requires independent verification and dual control.
2. **B.** An unexpected prompt may mean password compromise. Deny/report and escalate; never share codes.
3. **C.** Read right to left. `secure-check.invalid` is the controlling reserved domain; the brand-like labels to its left do not change ownership.
4. **C.** Voice and video are not sufficient identity proof. Initiate a callback and preserve financial controls.
5. **C.** Fast reporting lets defenders assess and protect others. Do not self-investigate or wait.
6. **B.** A QR code is an obscured link. Use the known portal.
7. **C.** Uncertainty is enough to report, and honest disclosure after interaction is urgent.
8. **B.** Polished language is cheap and scalable; inspect the action and verify the requester.
9. **A.** The requester must not control the verification route or contact data.
10. **C.** Evidence does not support treating awareness as a standalone barrier.

Employee quick reference

STOP — before you

- click or scan;
- open an unexpected attachment;

- sign in from a message;
- approve MFA;
- disclose a code, password, secret, or sensitive data;
- install software or grant remote control;
- change bank or payroll details; or
- send money or gift cards.

THINK

- Was this expected?
- What action am I being pushed to take?
- Is there urgency, secrecy, fear, authority, or reward?
- Could the account, voice, video, thread, or caller ID be compromised or fake?

VERIFY

- Start a new contact using the directory, known app, vendor master, contract, or in-person route.
- Never use contact information supplied in the suspicious request.
- Use dual control for payments and bank changes.
- No executive can waive verification inside the request.

REPORT

- Email: **[one-click button]**
- Teams/Slack: **[report menu / security channel]**
- SMS/call/QR/personal device: **[portal/address]**
- Clicked, entered credentials, approved MFA, ran software, disclosed data, or sent money: **call [urgent number] now**

> *Report even when unsure. False positives are welcome. If you made a mistake, say exactly what happened; fast reporting protects the company.*

Limitations and practical meaning

Threat reports do not measure one common denominator. IC3 counts voluntary complaints and adjusted reported loss. APWG aggregates reported attacks and contributing-member observations and notes collection effects. Microsoft reports detections in its ecosystem. Verizon analyzes incidents and breaches contributed by participating organizations. These sources illuminate different parts of the threat landscape; none estimates every phishing message or victim.

Attribution to AI is especially uncertain. A polished message is not proof of AI use, victims may not know whether synthetic media was involved, and an AI label can reflect many roles from drafting to voice cloning. This program therefore teaches a behavior that remains safe regardless of authorship.

Training studies also measure different outcomes: knowledge, intention, simulation clicks, reporting, or real incidents. Simulation performance is not equivalent to breach prevention. Large field studies provide important negative and modest results, often within one organization or sector. The correct practical interpretation is layered defense: make verification and reporting routine; continually test the full sociotechnical process; and engineer systems so one hurried decision does not become an account takeover or irreversible payment.

Visual assets

The following source figures are analytical, not decorative. Confirm the publication's reuse terms and obtain any required permission before redistributing them in later formats; linking here does not itself grant a reuse license.

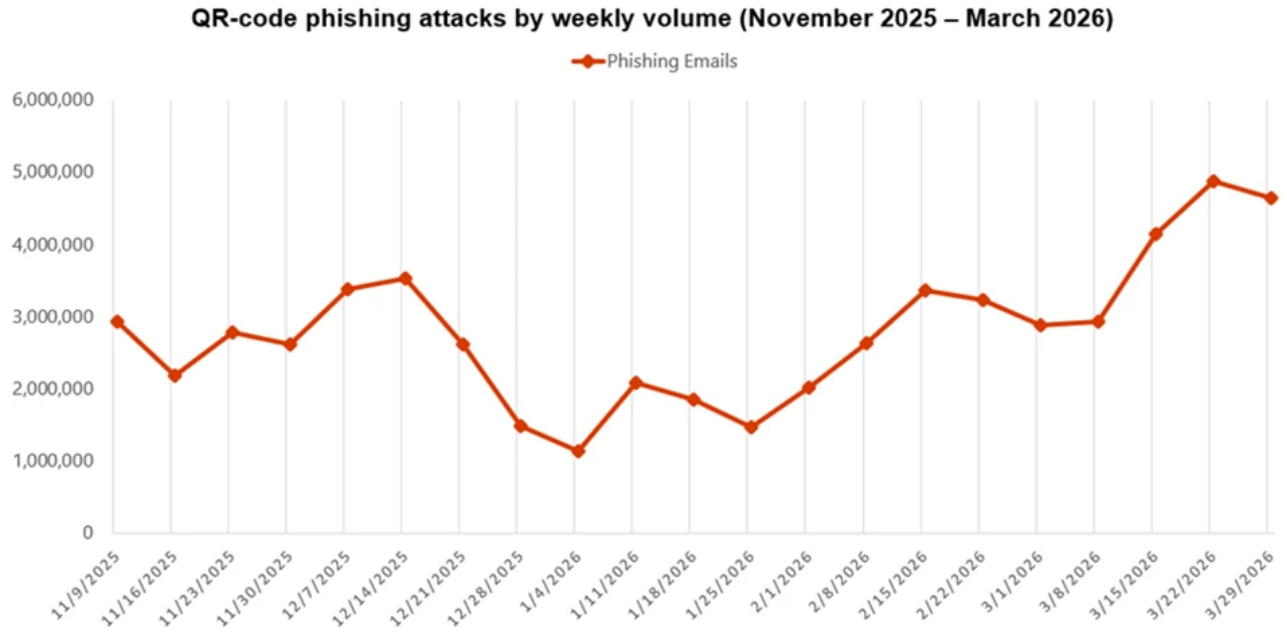

Figure — Line graph of weekly QR-code phishing volume from November 2025 through March 2026.

Credit: Microsoft Threat Intelligence and Microsoft Defender Security Research Team, Figure 3 in "Email threat landscape: Q1 2026 trends and insights".

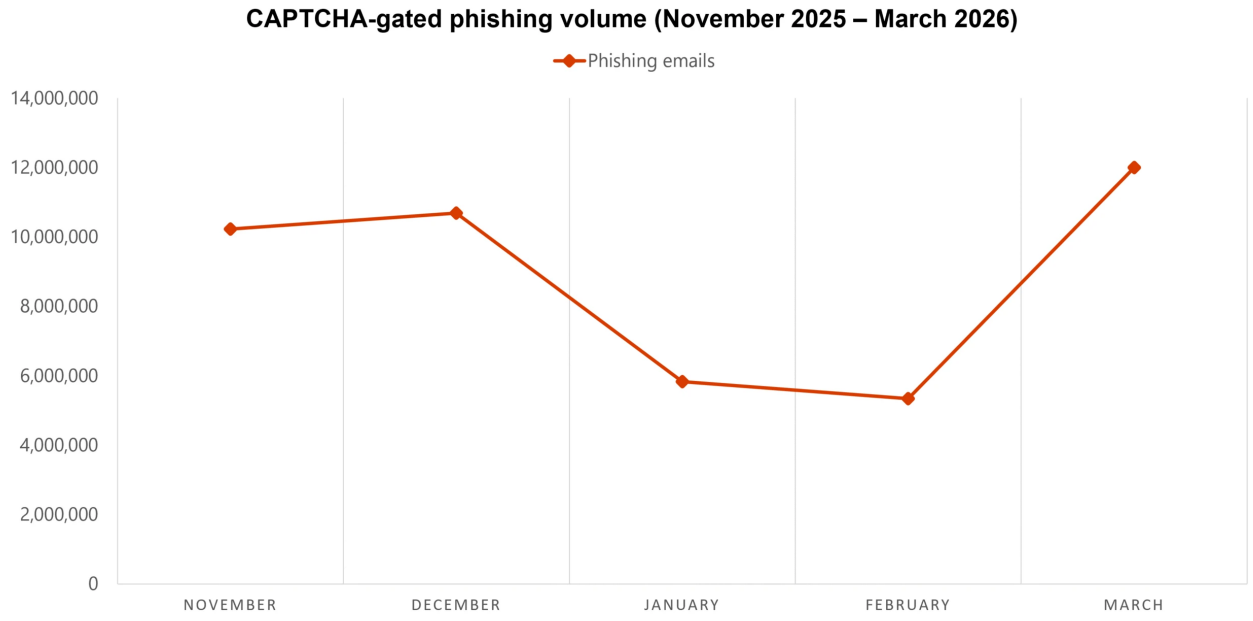

Figure — Line chart of CAPTCHA-gated phishing volume from November 2025 through March 2026.

Credit: Microsoft Threat Intelligence and Microsoft Defender Security Research Team, Figure 4 in "Email threat landscape: Q1 2026 trends and insights".

Sources

Claim supported	Evidence type	Source	Direct URL
2026 breach-entry and mobile social-engineering findings; human element remains important	Multi-contributor incident and breach analysis; vendor annual report	Verizon, *2026 Data Breach Investigations Report* and release	DBIR landing page and key findings release
2025 complaint counts and adjusted reported losses for phishing, BEC, and AI-related crime; BEC definition; rapid reporting and fund-recovery context	U.S. law-enforcement complaint data; voluntary reporting with stated definitions	FBI Internet Crime Complaint Center, *2025 IC3 Annual Report*	PDF
Q1 2026 reported phishing volume, telephone-fraud trend, social impersonation, and BEC cash-out observations	Industry clearinghouse report compiled from reports and member-contributed telemetry	APWG, *Phishing Activity Trends Report, 1st Quarter 2026*	PDF
Q1 2026 link, QR, attachment, CAPTCHA, and BEC trends in Microsoft detections	Large-scale vendor threat telemetry and threat-intelligence analysis	Microsoft Threat Intelligence and Defender Security Research Team, April 2026	Source page
Persistent Teams voice-phishing case involving IT-support impersonation	Incident-response case study	Microsoft Incident Response, March 2026	Source page
Teams abuse through chats, calls, meetings, external identities, device codes, files, and MFA social engineering	Vendor threat-intelligence analysis with observed actor cases	Microsoft Threat Intelligence, October 2025	Source page
Adversary-in-the-middle phishing can capture tokens and bypass non-phishing-resistant MFA	Technical campaign analysis	Microsoft Defender Security Research Team and Microsoft Threat Intelligence, May 2026	Source page
QR codes can conceal spoofed destinations; users should inspect previews and independently contact the organization	U.S. consumer-protection guidance based on scam reports	U.S. Federal Trade Commission, December 2023	Source page
Push bombing, MFA hierarchy, and migration to FIDO/WebAuthn; number matching as an interim measure	U.S. government technical guidance	Cybersecurity and Infrastructure Security Agency	Phishing-resistant MFA fact sheet and More than a Password
AI-generated voice impersonation and independent identity verification	U.S. law-enforcement public service announcement	FBI/IC3, May 2025	PSA
No relationship between annual-training recency and simulation failure; small embedded-training effect; engagement and lure-difficulty findings	Eight-month randomized controlled field experiment; >19,500 healthcare employees, ten campaigns	Ho et al., “Understanding the Efficacy of Phishing Training in Practice,” *IEEE Symposium on Security and Privacy*, 2025	Open paper
Repeated simulation click rates declined, but targeted mandatory training did not substantially improve the high-risk group	Longitudinal observational healthcare study; 5,416 employees receiving 20 campaigns	Gordon et al., “Evaluation of a mandatory phishing training program for high-risk employees,” *Journal of the American Medical Informatics Association*, 2019	Open article
Click rates are confounded by lure difficulty and work-context alignment; need difficulty-aware interpretation	Peer-reviewed operational-data study and scale development	Steves, Greene, and Theofanos, “Categorizing Human Phishing Detection Difficulty,” *Journal of Cybersecurity*, 2020	NIST record and DOI
How to apply the Phish Scale consistently and tailor it to an organization	Government practitioner guide grounded in empirical research	NIST Technical Note 2276, *NIST Phish Scale User Guide*, 2023	PDF
Work context changes how employees interpret phishing cues	4.5-year workplace-situated mixed-method study in a small U.S. government research unit	Greene et al., “User Context: An Explanatory Variable in Phishing Susceptibility,” 2018	NIST record
Large-enterprise evaluation requires correction for bias and careful methodology	Analysis of 115,080 simulated emails to 19,180 participants at one organization	Siadati et al., “Measuring the Effectiveness of Embedded Phishing Exercises,” USENIX CSET 2017	USENIX paper page
Motivation, communication, feedback, and social context affect employee engagement in phishing interventions	Peer-reviewed usable-security study	Chen et al., “What Motivates and Discourages Employees in Phishing Interventions,” SOUPS 2024	Open paper
Building an organization-wide, behavior- and culture-focused cybersecurity awareness and training program	U.S. government standard/guidance	NIST SP 800-50 Rev. 1, *Building a Cybersecurity and Privacy Learning Program*, 2024	NIST record