Servicio Nacional del Consumidor

FACULTAD ECONOMÍA Y NEGOCIOS UNIVERSIDAD DE CHILE

Fraude en Medios de Pago en Chile
Estudio Experimental: ¿Cómo reforzar las Campañas de prevención de las estafas en línea?

Subdirección de Consumo Financiero
Coordinación de Economía del Comportamiento

Universidad de Chile
Facultad de Economía y Negocios
Departamento de Administración

Julio de 2025

Resumen ejecutivo

Según el *World Economic Forum*, la ciberseguridad es uno de los riesgos globales críticos a corto plazo (WEF, 2024). El uso de tecnologías avanzadas como la inteligencia artificial y el aprendizaje automático ha incrementado la sofisticación de los ataques fraudulentos, transformando la seguridad en un desafío constante para individuos, organizaciones y gobiernos (Edwards, 2024). En Chile, las estadísticas sobre reclamaciones de fraudes en medios de pago digital, presentadas por los consumidores a las entidades financieras, muestran un notable incremento en los últimos años. Durante el segundo semestre de 2023, se alcanzó el nivel más alto de reclamaciones. Los montos reclamados en 2023 en las instituciones bancarias más que se duplicaron respecto del año anterior, alcanzando aproximadamente $243 mil millones. En tanto, en el 2024 alcanzaron $275 mil millones.

En este contexto, la Ley N° 21.673, publicada en el Diario Oficial el 30 de mayo de 2024, reformuló el marco regulatorio relacionado a la limitación de la responsabilidad de los usuarios de medios de pago con ocasión del fraude, contenido en la Ley N° 20.009. Respecto de la prevención del fraude, la normativa indica que los **usuarios deberán informarse y adoptar todas las medidas necesarias para prevenir** el uso indebido, el fraude u otros riesgos afines a la utilización de los medios de pago y los mecanismos de autenticación asociados, y que las **entidades reguladas tienen el deber de proveer información periódica, clara, accesible y actualizada sobre las medidas de seguridad y las instrucciones para un uso seguro**, fomentando prácticas responsables en la gestión de los medios de pago.

En este sentido, la reforma destaca el papel fundamental del factor humano en la seguridad informática. Aunque la tecnología es esencial, por sí sola no es suficiente para prevenir los fraudes, por lo que fomentar comportamientos a favor de la autoprotección de los consumidores, a través de diversas campañas de ciberseguridad, juega un rol fundamental (Liang & Xue, 2010; Ng Boon-Yuen, et al. 2009; Rhee, et al. 2009).

Las campañas de ciberseguridad buscan, principalmente, sensibilizar a los consumidores sobre las amenazas, fomentar hábitos seguros y educar para cambiar percepciones sobre la seguridad digital. Sin embargo, no todas las entidades financieras ofrecen campañas integrales contra el fraude en medios de pago en línea, o si lo hacen, no cubren todos los tipos de fraude existentes. A menudo, la información disponible es incompleta y de difícil acceso (SERNAC, 2025a). Por ello, la aplicación de las ciencias conductuales en estas campañas es fundamental. Permiten diseñar estrategias que promuevan eficazmente la adopción de hábitos preventivos, mitigando así los riesgos de fraude (Anderson et al., 2010; Acquisti et al., 2015).

De acuerdo con la **Teoría de la Acción Razonada** y la **Teoría del Comportamiento Planificado,** las personas son más propensas a adoptar comportamientos seguros, como seguir recomendaciones de ciberseguridad, cuando se influye en ellas a través de normas subjetivas (la presión percibida por el entorno), el control conductual percibido (la creencia en la propia capacidad y recursos para actuar), y, crucialmente, un cambio de actitud (la valoración personal positiva del comportamiento) (Fishbein y Ajzen, 1975; Ajzen, 1991; Ajzen y Madden, 1986).

Las campañas de concientización son herramientas de comunicación diseñadas para influir y persuadir a las personas, buscando generar un cambio de comportamiento observable. La influencia se refiere a cualquier modificación en los estados mentales o acciones de las personas provocada por un estímulo externo (Cialdini, 2009; Perloff, 2003; Pratkanis & Aronson, 1994), mientras que la persuasión es un tipo específico de influencia que busca cambiar la actitud hacia un objeto, persona o tema (Petty & Cacioppo, 1986).

El estudio científico de la persuasión ha evolucionado con modelos influyentes. **El Modelo de Comunicación y Persuasión de Hovland/Yale** (Hovland et al., 1953; Hovland & Janis, 1959) propuso que el cambio de actitud ocurre mediante un proceso de aprendizaje que implica atención, comprensión y aceptación de la información. Ampliando esto, el **Modelo de Procesamiento de la Información** de William McGuire (1969; 1985) describe la persuasión como una secuencia de seis etapas: exposición, atención, comprensión, aceptación, retención y, finalmente, la conducta. Este modelo subraya que el éxito persuasivo depende de completar cada uno de estos pasos.

En este contexto, el estudio realizado por Sernac tuvo como **objetivo identificar elementos comunicacionales que mejoren la efectividad de las campañas contra el Phishing**. El phishing es una técnica de fraude cibernético diseñada para engañar a las personas y que revelen información sensible, como credenciales o datos financieros. Se ejecuta principalmente a través de correos electrónicos, mensajes o sitios web fraudulentos (Ferreira et al., 2015). A diferencia de los ataques que explotan sólo vulnerabilidades técnicas, el phishing se basa en la manipulación psicológica de sus víctimas, utilizando principios de ingeniería social para inducirlas a actuar en beneficio del atacante (Jari, 2022).

La efectividad de una campaña, en este estudio, se mide considerando las siguientes variables:

- **Probabilidad de diferenciación correcta**: La capacidad del usuario para distinguir correctamente entre correos electrónicos fraudulentos y legítimos.
- **Número de elementos identificados:** La cantidad de señales de phishing que los usuarios reconocen y utilizan para clasificar los correos.
- **Actitud hacia la autoprotección:** La mejora en la percepción y valoración de las personas sobre la importancia de protegerse contra el phishing.
- **Intención de seguir recomendaciones**: El aumento en la disposición de los usuarios para aplicar las medidas sugeridas para mitigar los riesgos de fraude.
- Además, se evalúan **variables relacionadas con el canal de persuasión**, tales como la atención y comprensión del mensaje, la fuerza argumental percibida, la efectividad general percibida de la campaña, el agrado general hacia el contenido, la gravedad percibida de la amenaza y la autoeficacia percibida por el usuario para enfrentar estos riesgos.

Para potenciar las campañas estándar de detección de correos fraudulentos, se evaluó el impacto de tres factores comunicacionales específicos:

Mensajes Motivacionales: Estos mensajes refuerzan la autoeficacia del usuario y subrayan la gravedad del phishing. Un estudio previo del SERNAC (marzo-abril de 2024) ya destacó la autoeficacia como el factor más determinante para fomentar la autoprotección, seguido de la percepción de la gravedad del fraude (Sernac, 2025b).

Explicación de Técnicas de Ingeniería Social: Se proporciona una breve descripción de los métodos de manipulación usados en correos fraudulentos, lo que facilitaría su identificación por parte de los usuarios.

Enfoque Lúdico ("Juego"): Este consiste en un test donde los participantes evalúan la legitimidad de correos. Permite aplicar las recomendaciones recibidas previamente, contribuyendo a un proceso de persuasión más efectivo y fomentando una mayor inmersión en la narrativa de la campaña.

Para entender como los "Juegos" potencian la persuasión, se utiliza **el Modelo de Probabilidad de Elaboración Extendida (EELM)** de Slater y Rouner (2002), el cual explica cómo y por qué se imitan los comportamientos observados. Este enfoque sostiene que el aprendizaje se refuerza mediante la observación. Sin embargo, la simple exposición no garantiza la adopción de una conducta; la motivación juega un papel crucial (Slater y Rouner, 2002; Moyer-Gusé, 2008). El EELM analiza cómo diversos factores influyen en la motivación y, por ende, en la adopción de actitudes o comportamientos. Por ejemplo, la resistencia a los mensajes puede obstaculizar esta motivación. No obstante, al presentar una narrativa persuasiva a través de un juego, la inmersión del individuo aumenta, reduciendo su criticidad y resistencia al mensaje (Shrum, 2004; Slater y Rouner, 2002).

Este proceso de persuasión se facilita por constructos como la identificación, la homofilia y la transportación:

- **Identificación:** Implica la inmersión del jugador en el entorno del personaje, permitiendo una comprensión cognitiva y emocional profunda. Esto genera respuestas emotivas, como la empatía, que pueden motivar al individuo a alinear sus objetivos con los del personaje (Cohen, 2001).
- **Homofilia:** Se refiere a la similitud percibida entre el individuo y un personaje clave, basada en la creencia de compartir conocimientos, emociones o metas. A diferencia de la identificación, no es una experiencia indirecta, como vivir las emociones de otro, sino una conexión por atributos comunes, como preferencias o aversiones.
- **Transportación:** Se enfoca en la absorción total del individuo en la narrativa, donde la persona se "pierde" completamente en la historia (Green y Brock, 2000). En este estado de inmersión, el enfoque ya no recae en la elaboración o el procesamiento consciente del contenido narrativo, sino en una inmersión total en la experiencia (Gerrig, 1993).

Respecto de la Metodología utilizada, para evaluar la efectividad de las campañas de ciberseguridad, se implementó un **Experimento Controlado Aleatorizado (RCT)** con la participación de 6.044 consumidores, cuyo levantamiento de información se realizó entre noviembre y diciembre del 2024. Este diseño factorial permitió analizar los impactos tanto a nivel de tratamientos específicos como de factores individuales.

Los RCTs son la metodología experimental más rigurosa para determinar el impacto causal de una intervención. Su fortaleza radica en la aleatorización de los participantes en grupos de tratamiento y control. Esto asegura que los grupos sean estadísticamente equivalentes en todas sus características (observables y no observables), eliminando sesgos y permitiendo atribuir cualquier diferencia observada directamente a la intervención (Gertler et al., 2016; Duflo et al., 2006). De esta forma, se puede estimar el "contrafactual" —lo que habría ocurrido sin la intervención— comparando los resultados del grupo tratado con los del grupo de control.

Para medir el impacto, se utilizaron dos enfoques estadísticos clave:

Modelo Basado en Tratamientos (Cell Means Model): Este modelo estima y compara directamente las medias de cada grupo experimental. Es útil para evaluar el impacto de diferentes condiciones tratándolas como unidades independientes, comparando los promedios de un indicador específico entre el grupo tratado y el de control. Si las variables son continuas y el tamaño muestral es grande, se aplica la prueba t de Student.

Análisis de Varianza Factorial (ANOVA Factorial): Esta técnica permite analizar el efecto de múltiples factores categóricos sobre una variable dependiente continua, evaluando tanto sus efectos individuales (principales) como sus interacciones. El ANOVA factorial es crucial para identificar cómo los diferentes elementos comunicacionales (factores) influyen en la efectividad y cómo se combinan entre sí (Field, 2018;

Montgomery, 2019). La inferencia se basa en la prueba F, que determina si los factores explican una variabilidad significativa en la variable dependiente.

Los resultados principales son los siguientes:

1. El "Juego" es el elemento con mayor impacto en los indicadores principales: la actitud hacia las medidas de detección de estafas en línea, y la intención de protegerse.
2. La "Explicación" mejora la aceptación general de la campaña y la percepción de autoeficacia, aunque su efecto suele ser más notorio en ausencia del "Juego" o en combinación con la "Motivación".
3. Finalmente, la "Motivación" por sí sola no mostró resultados significativos, sin embargo, cuando se combina con la "Explicación", se observan mejoras significativas en su influencia.

Lo anterior también es válido para los resultados que se presentan a continuación, los impactos sobre los indicadores intermedios, es decir, sobre las variables del canal de persuasión (atención, comprensión, fuerza argumental percibida, efectividad percibida, gusto general, gravedad percibida, autoeficacia percibida):

- Se encuentra que el factor *juego* actúa a través de todos los canales propuestos, confirmando su relevancia.
- La *Explicación* actúa aumentando la percepción de *Autoeficacia, Efectividad percibida, Gusto general por la campaña y Fuerza argumental percibida.*
- La interacción entre los tres factores aumenta la gravedad percibida y la fuerza argumental percibida.

Por otro lado, el análisis de heterogeneidad, realizado desde la perspectiva de los tratamientos, encontró los siguientes resultados:

- No se encuentran efectos heterogéneos por género sobre las variables de *actitud* e *intención*, lo que indica que los efectos de los tratamientos son similares tanto para hombres como para mujeres
- El análisis de heterogeneidad por edad revela que los *adultos-jóvenes* (<45 años) tienden a tener una *actitud* menos favorable hacia las medidas de seguridad y una menor *intención* de protegerse. No obstante, el tratamiento que combina *explicación y juego* (T3') impacta positivamente en su *actitud*, mientras que la interacción entre explicación, *motivación* y *juego* (T4') influye significativamente en su *intención* de protección.
- Además, la *intención* de protección de los *adultos-jóvenes* (<45 años) es afectada positivamente por el tratamiento *juego* (T1') y el tratamiento que combina *explicación*, *motivación* y *adultos jóvenes* (T8').
- En resumen, estos hallazgos sugieren que, aunque los *adultos jóvenes* presentan menor disposición inicial hacia la seguridad, las estrategias comunicacionales que combinan *juego*, *motivación* y *explicación* pueden mejorar tanto su *actitud* como su *intención* de protegerse.

Asimismo, es importante destacar que la inclusión de la Explicación de Técnicas de Ingeniería Social en las campañas tuvo otro efecto positivo: los participantes no solo identificaron más elementos fraudulentos en los correos, sino que también se tomaron más tiempo para revisarlos. Este resultado es especialmente alentador, ya que las estafas de phishing a menudo explotan las respuestas impulsivas de las víctimas. El mayor tiempo de revisión sugiere un incremento en la precaución y el análisis crítico por parte de los participantes, lo que los hace menos vulnerables.

Por último, se destaca el resultado de la pregunta del estudio: "¿A través de qué medio preferiría acceder a las campañas de seguridad provenientes de organismos públicos e instituciones financieras?". Se encontró que:

- Los medios preferidos por los participantes fueron el email, el video, el sitio web del SERNAC y el sitio web de las instituciones financieras.
- Estas preferencias son las mismas para hombres y mujeres.
- Al diferenciar por grupo etario, se encuentra que el grupo de 18-29 años prioriza el video, mientras que el grupo de 60 años o más prefiere el email como principal medio de acceso a las campañas.

Los resultados de este estudio pueden ser usados para mejorar las campañas comunicacionales de SERNAC, al considerar los elementos identificados en su diseño (como en los mensajes comunicacionales utilizados en el experimento), o bien, para hacer recomendaciones de buenas prácticas a la industria.

Recomendaciones de política

Dado los resultados mencionados, se proponen las siguientes recomendaciones claves:

1. Optimizar las Campañas de Concientización

Es crucial reforzar las campañas de concientización existentes con nuevos recursos comunicacionales que impulsen la adopción proactiva de medidas de seguridad. Específicamente:

- **Implementar juegos serios**: El uso de juegos interactivos es altamente recomendable en sus diversas modalidades. Incorporan mecanismos que favorecen el aprendizaje, la transformación de actitudes y la adopción de nuevos comportamientos.

- **Explicar las tácticas de ingeniería social**: En las campañas dirigidas contra el *phishing*, dadas sus sofisticadas técnicas de manipulación, es fundamental explicar detalladamente las tácticas de ingeniería social más comunes. Esto permitirá a las personas reconocer la manipulación y actuar con mayor conciencia.

- **Incorporar mensajes motivacionales:** Sugerimos incluir frases motivacionales que refuercen la autoeficacia del usuario y la gravedad del

phishing. Estos mensajes deben complementar otros elementos comunicacionales de la campaña.

2. Diversificar y Personalizar la Difusión

Para maximizar el alcance y la efectividad, es importante utilizar diversos medios para difundir la información, como el correo electrónico, videos, el sitio web del SERNAC y los portales de instituciones financieras. Además, se debe personalizar la entrega de mensajes según el perfil sociodemográfico de cada consumidor, asegurando que el contenido sea relevante y resuene con cada segmento.

3. Evaluar Continuamente las Campañas

Finalmente, es esencial evaluar periódicamente la efectividad de las campañas. Esto se puede lograr a través de juegos interactivos y encuestas de percepción. Esta evaluación continua permitirá asegurar el cumplimiento de los estándares mínimos y realizar los ajustes necesarios cuando los resultados no alcancen las expectativas.

El diagnóstico presentado en este informe busca contribuir a que los distintos componentes del ecosistema financiero –i.e. emisores, usuarios y supervisores- mitiguen el riesgo de fraude al consumidor.