

Switch\_

# Human-centred Security meets AI

## How to navigate new threats

Cornelia Puhze

Swiss Cyber Storm, 22 October 2024

1. Meeting compliance is enough to mitigate human risk.
2. Our approach to Security Awareness Training is science-based.
3. Behaviour change is complex to achieve.
4. Using fear to convince people to behave securely is effective.
5. Phishing simulations are effective and should be part of every SAT programme.

# Switch

**NREN**

-

National Research  
and Education  
Network

**Registry**

-

for .ch/.li ccTLDs

**Security Advocacy**

**Events & Trainings for Practitioners**

**Security Awareness Adventures**

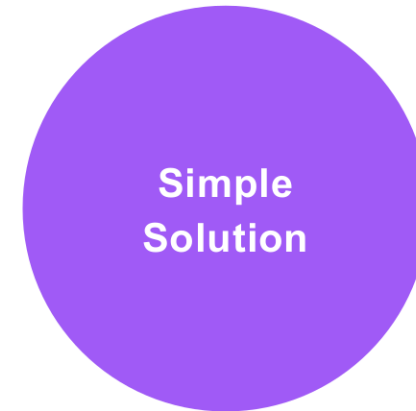
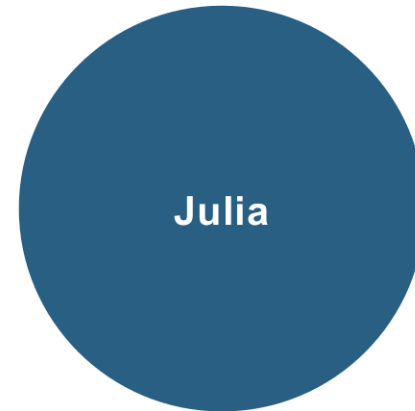


**Competence Centre  
for Security Awareness**  
Switch, Zurich

→[awareness@switch.ch](mailto:awareness@switch.ch)

→<https://swit.ch/security-awareness>

# Agenda



# What's the impact of AI?

# Human hacking: targeted, cheap, much harder to detect

Figure 33: Time series of major incidents observed by ENISA (July 2023-June 2024)



## Protecting against the people problem

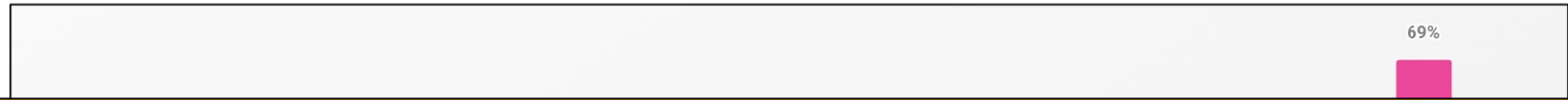
Proofpoint 2024 Voice of the CISO report

To mitigate this area of human vulnerability, many CISOs are turning to AI-powered technology. Of those surveyed, **87%** are looking to deploy such tools to protect against human error and block advanced human-centric cyber threats.

© 2023/07

Base: US, Canada, UK, Germany, Australia, New Zealand, and India. Total number of participants: 7012 (age 18+).  
Dates conducted: March 6, 2024 - April 22, 2024.

# What's real?



**The average accuracy is around  
50% for untrained observers,  
60% for trained observers with unlimited time.**



*Cyber Security!*

**Cyber Security!**

*Cyber Security!*



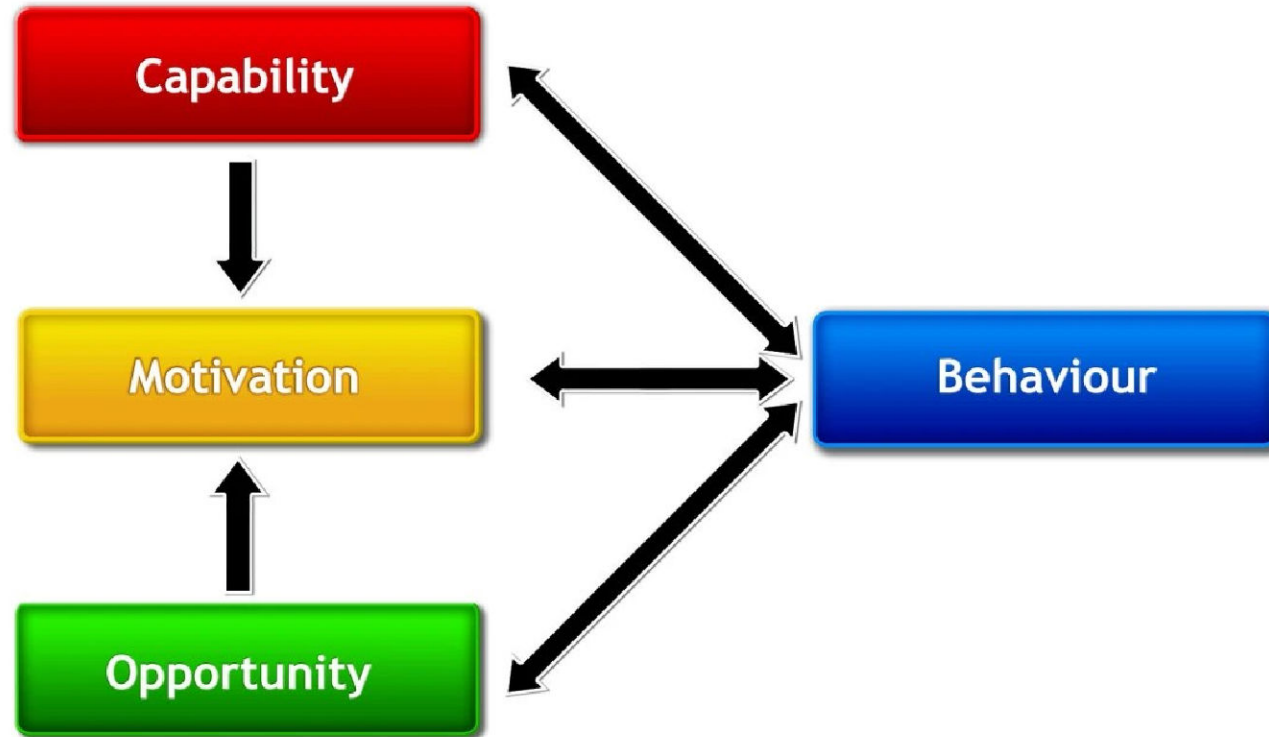
**Security managers typically only consider lack of knowledge ... Thus, their current efforts in „security education“ consist of repeating all policies and rules to everyone. This is the equivalent of shouting louder at someone who does not understand your language; we need a smarter, targeted approach if we want to meaningfully change behavior...**

(Beris et al., 2015)

# The path to behaviour change



# The behaviour change wheel: COM-B



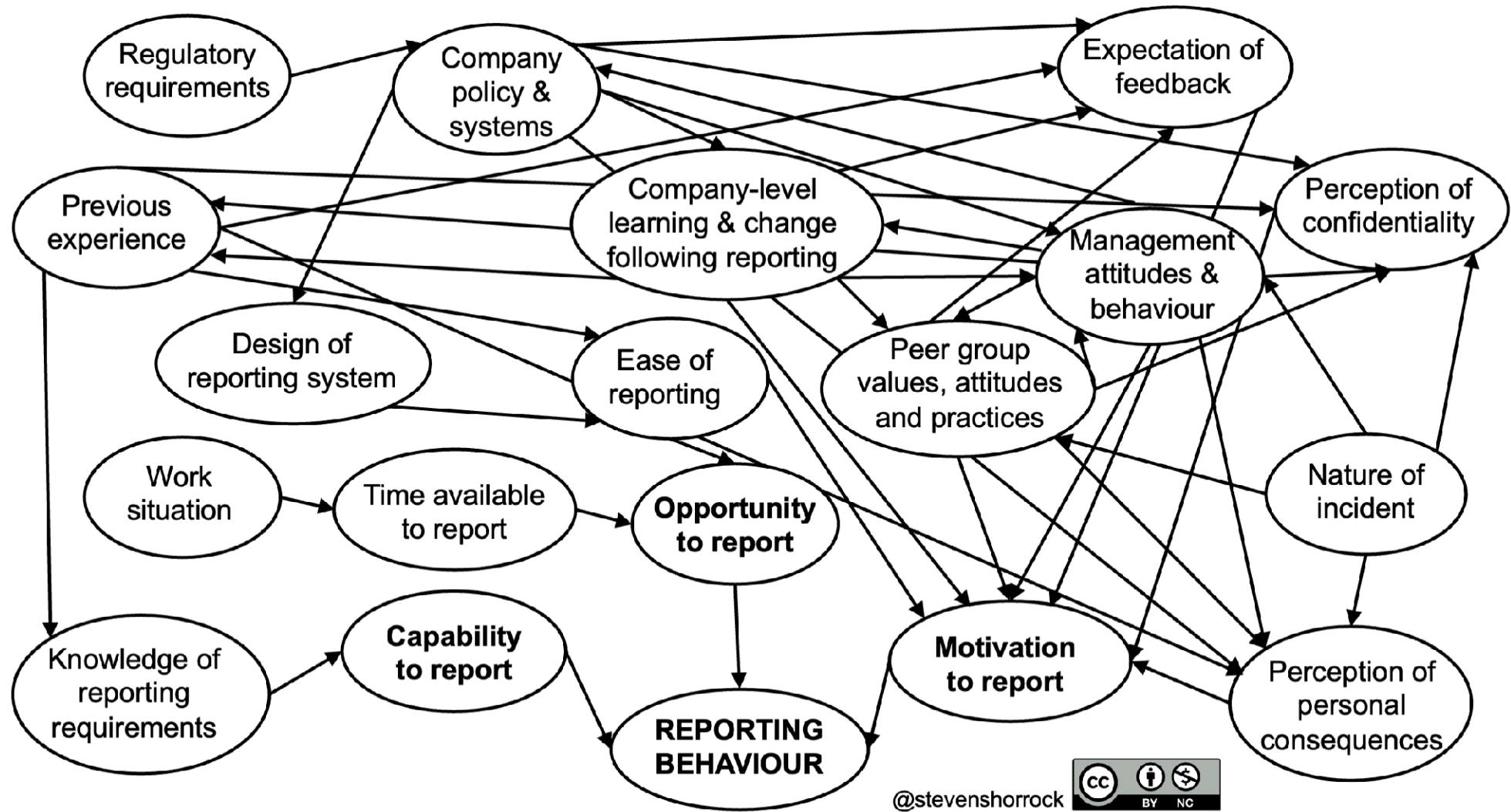
Michie, S., van Stralen, M.M. & West, R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Sci* 6, 42 (2011). <https://doi.org/10.1186/1748-5908-6-42>

# Modeling influences on behaviour



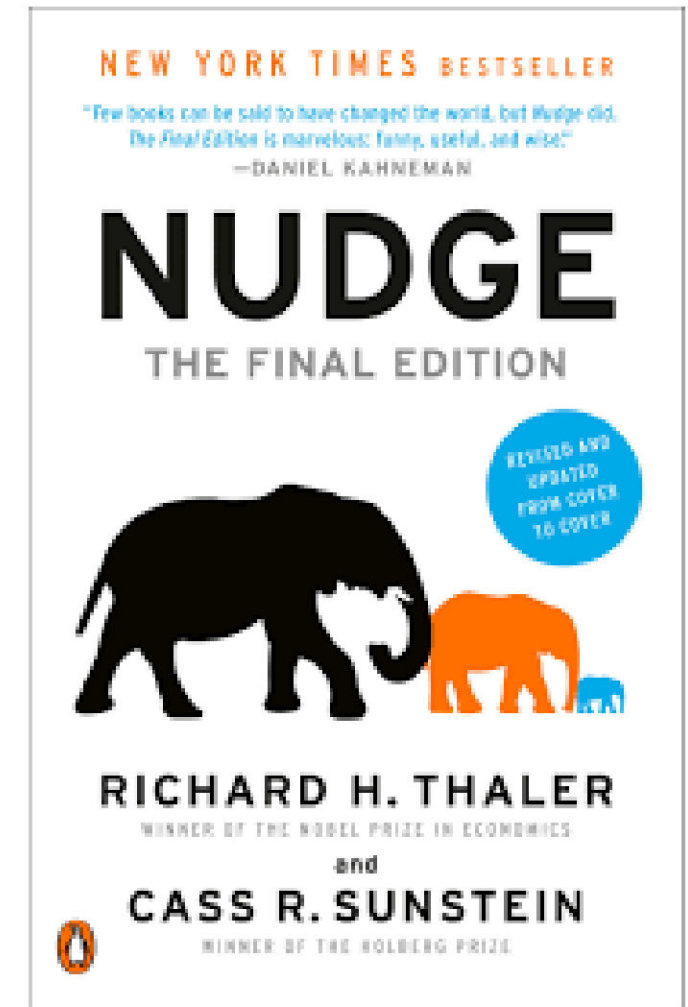
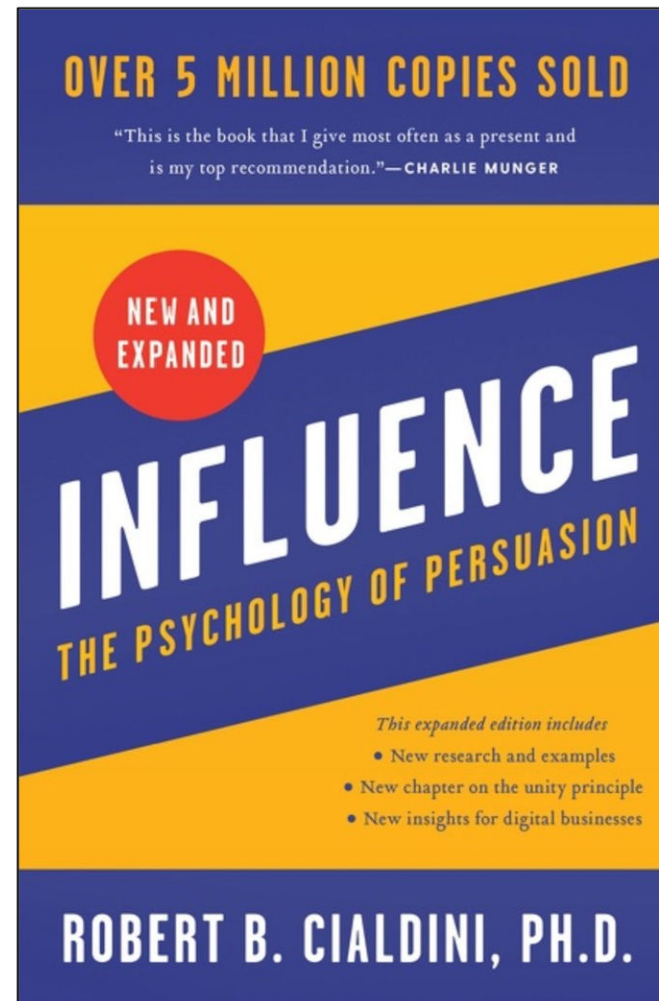
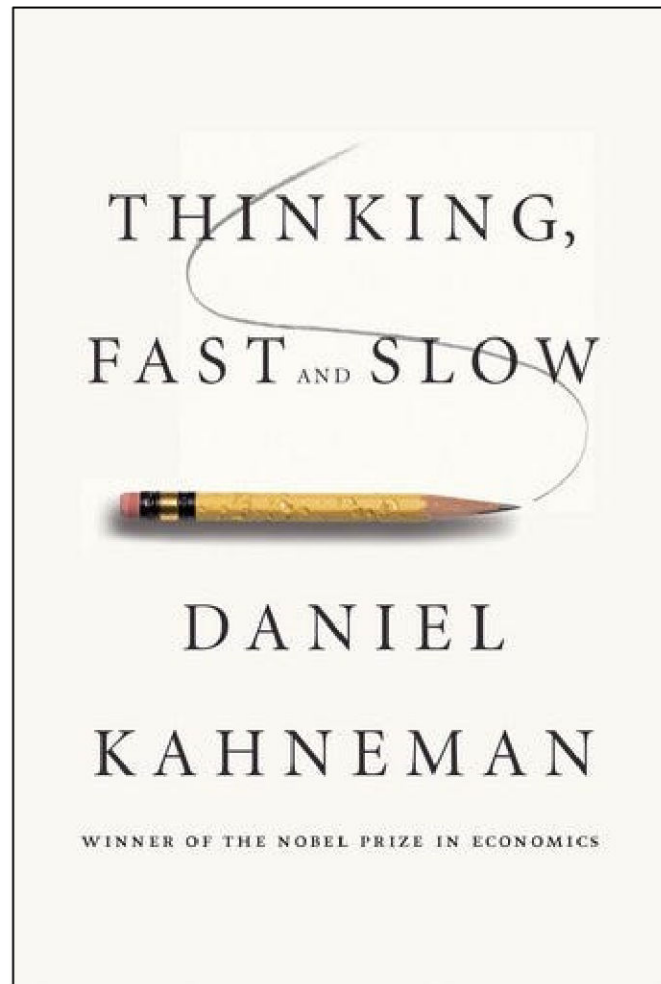
Shorrock, S. (2023, November 17). "Why aren't they reporting incidents?" Influences on reporting behaviour. *Humanistic Systems*. [humanisticsystems.com/2023/11/14/why-arent-they-reporting-incidents-influences-on-reporting-behaviour/](https://humanisticsystems.com/2023/11/14/why-arent-they-reporting-incidents-influences-on-reporting-behaviour/)

# “Why aren’t they reporting incidents?”

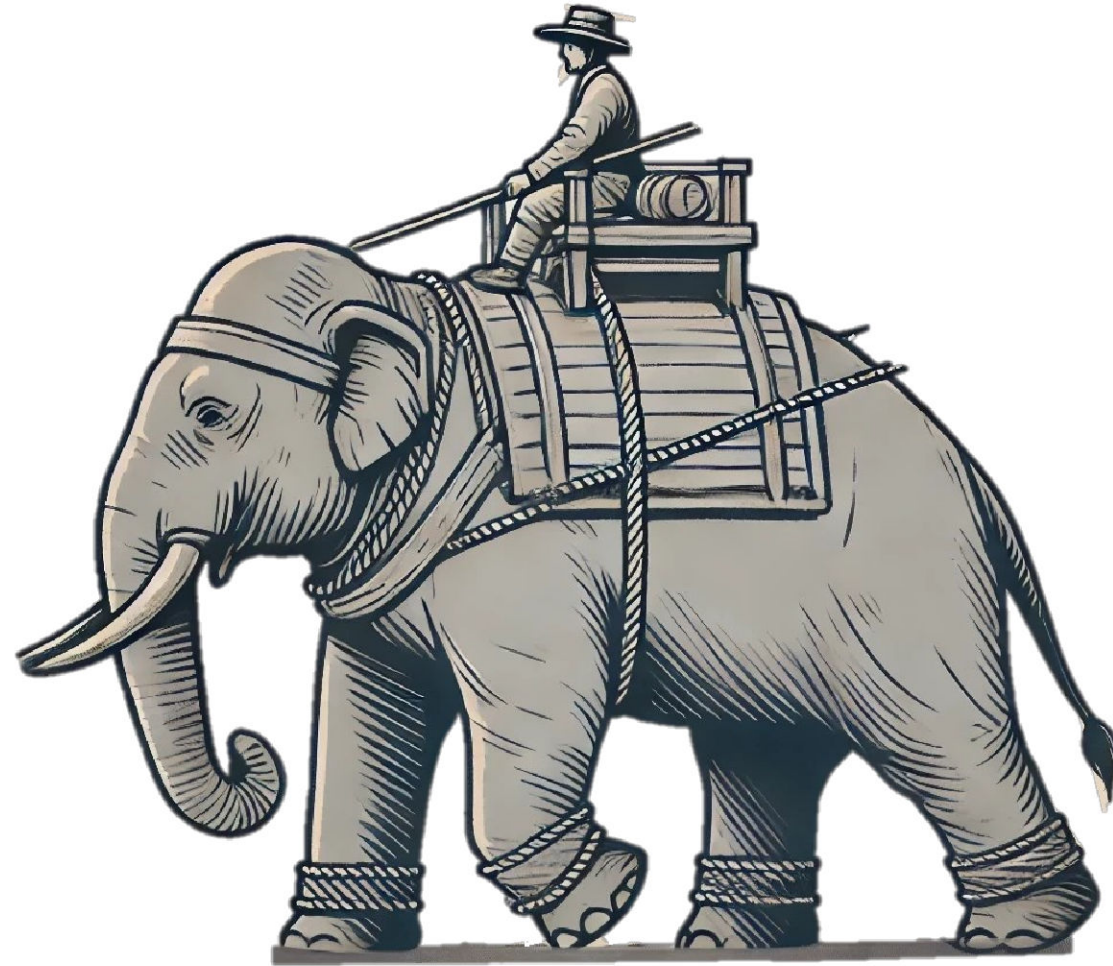


@stevenshorrock

# Human OS



# The Elephant and Rider Metaphor





# Julia

- Reputation & Trust
- Learning & Training
- Gut feeling



# USERS ARE NOT STUPID:

## Six Cybersecurity Pitfalls Overturned



For more information, visit  
<https://csrc.nist.gov/usable-cybersecurity>

### PITFALLS & MISCONCEPTIONS

#### 1. Assuming users are stupid

Thinking users are stupid or hopeless creates an “us vs. them” situation that puts security professionals at odds and reduces users’ confidence in their own ability to make cybersecurity decisions.

#### 2. Not tailoring cybersecurity communication

When communicating security information, security professionals often fail to tailor the message to the audience.

# Google Security Blog

The latest news and insights from Google on security and safety on the Internet

## On Fire Drills and Phishing Tests

May 22, 2024

Matt Linton, Chaos Specialist



## An investigation of phishing awareness and education over time: When and how to best remind users

Benjamin Reinheimer, Lukas Aldag, Peter Mayer, Mattia Mossano, and Reyhan Duezguen, *SECUSO - Security, Usability, Society, Karlsruhe Institute of Technology*; Bettina Lofthouse, *Landesamt für Geoinformation und Landesvermessung Niedersachsen*; Tatiana von Landesberger, *Interactive Graphics Systems Group, Technische Universität Darmstadt*; Melanie Volkamer, *SECUSO - Security, Usability, Society, Karlsruhe Institute of Technology*

<https://www.usenix.org/conference/soups2020/presentation/reinheimer>



## Phishing-Kampagnen zur Mitarbeiter-Awareness

Analyse aus verschiedenen Blickwinkeln: Security, Recht und Faktor Mensch

12.05.2020

## Phishing in Organizations: Findings from a Large-Scale and Long-Term Study

Daniele Lain, Kari Kostiainen, and Srdjan Čapkun  
*Department of Computer Science*  
*ETH Zurich, Switzerland*  
{daniele.lain, kari.kostiainen, srdjan.capkun} @inf.ethz.ch

# Phishing in Organizations: Findings from a Large-Scale and Long-Term Study

Daniele Lain, Kari Kostiainen, and Srdjan Čapkun  
*Department of Computer Science*  
*ETH Zurich, Switzerland*  
{daniele.lain, kari.kostiainen, srdjan.capkun} @inf.ethz.ch

2021, 15 months, 14,000 study participants

«...**embedded training** during **simulated phishing exercises**, does make employees **NOT more resilient to phishing**, but (... it) can make employees even **more susceptible to phishing**.»

# Content, Nudges and Incentives: A Study on the Effectiveness and Perception of Embedded Phishing Training

Daniele Lain  
ETH Zurich  
Department of Computer Science  
Zurich, Switzerland

Tarek Jost  
ETH Zurich  
Department of Computer Science  
Zurich, Switzerland

Sinisa Matetic  
ETH Zurich  
Department of Computer Science  
Zurich, Switzerland

Kari Kostiainen  
ETH Zurich  
Department of Computer Science  
Zurich, Switzerland

Srdjan Capkun  
ETH Zurich  
Department of Computer Science  
Zurich, Switzerland

2024, <https://arxiv.org/pdf/2409.01378>

«Phishing is an **attention problem**, rather than a knowledge one, even for the most susceptible employees, and thus enforcing training does not help.»

# Gut feeling?

- Rule-based training: 42% fewer clicks
- Mindfulness training: 151% fewer clicks

## Training to Mitigate Phishing Attacks Using Mindfulness Techniques

Matthew L. Jensen, Michael Dinger, Ryan T. Wright & Jason Bennett Thatcher

To cite this article: Matthew L. Jensen, Michael Dinger, Ryan T. Wright & Jason Bennett Thatcher (2017) Training to Mitigate Phishing Attacks Using Mindfulness Techniques, Journal of Management Information Systems, 34:2, 597-626, DOI: [10.1080/07421222.2017.1334499](https://doi.org/10.1080/07421222.2017.1334499)

To link to this article: <https://doi.org/10.1080/07421222.2017.1334499>

SLOW DOWN ....  
& FROWN



When reading this email

computer  
**FRAUD & SECURITY**

[Journal Home](#) | [Current Issue](#) | [Archive](#) ▾ | [Subscribe](#) | [About](#)

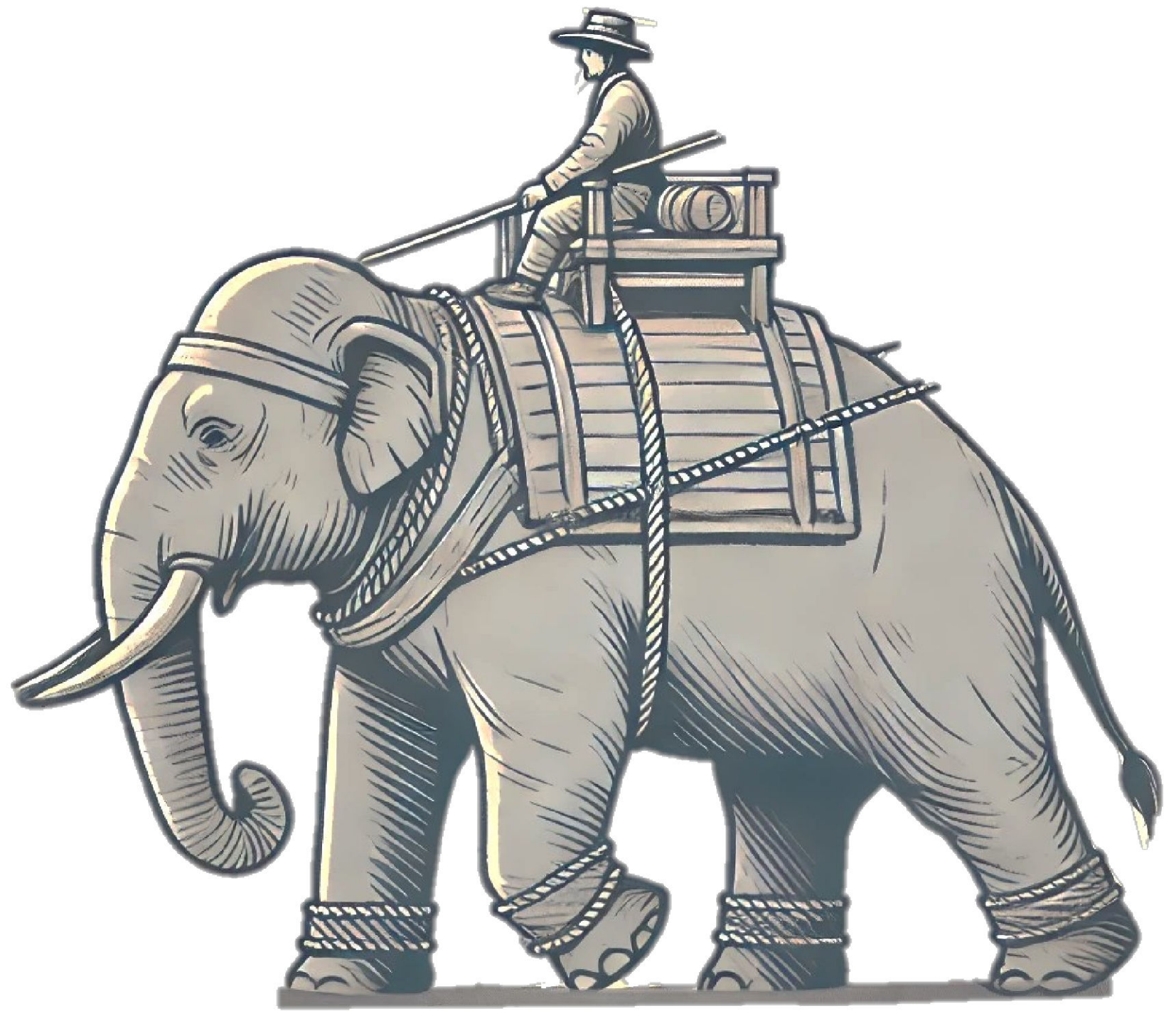
Current Issue

Computer Fraud & Security, Vol. 2022, No. 10 • Features

## Slow down and frown to improve phishing detection

Patricia Nevin, Karen Renaud, George Finney

Published Online: 29 Oct 2022 | [https://doi.org/10.12968/S1361-3723\(22\)70593-3](https://doi.org/10.12968/S1361-3723(22)70593-3)



1. Meeting compliance is enough to mitigate human risk.
2. Our approach to Security Awareness Training is science-based.
3. Behaviour change is complex to achieve.
4. Using fear to convince people to behave securely is effective.
5. Phishing simulations are effective and should be part of every SAT programme.



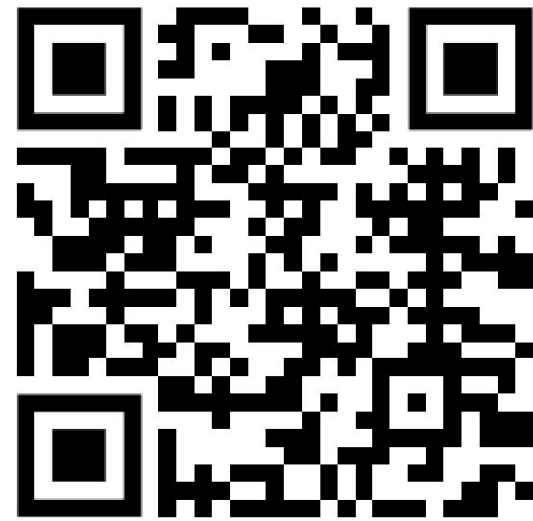
# Swiss Security Awareness Day

October 24th, 2024  
Zentrum Paul Klee  
Bern



# In for some action?

# Security Awareness Adventures



# CAS Cyber Risk Awareness (Spring 2025)



**Switch**

**cyren<sup>zh</sup>**

Zürich University  
of Applied Sciences  
**zhaw** School of  
Management and Law

# Contact

[awareness@switch.ch](mailto:awareness@switch.ch)

# Fear Appeals

## Cyber Security Fear Appeals: Unexpectedly Complicated

Karen Renaud  
University of Abertay, Dundee, UK  
k.renaud@abertay.ac.uk

Marc Dupuis  
University of Washington, USA  
marcjd@uw.edu

### ABSTRACT

Cyber security researchers are starting to experiment with fear appeals, with a wide variety of designs and reported efficaciousness.

recommended action, and, second, that eliciting the fear emotion by highlighting unpleasant consequences is likely to make them care.

*Why not just tell people what to do? The problem is that knowl*

NSPW '19, September 23–26, 2019, San Carlos, Costa Rica

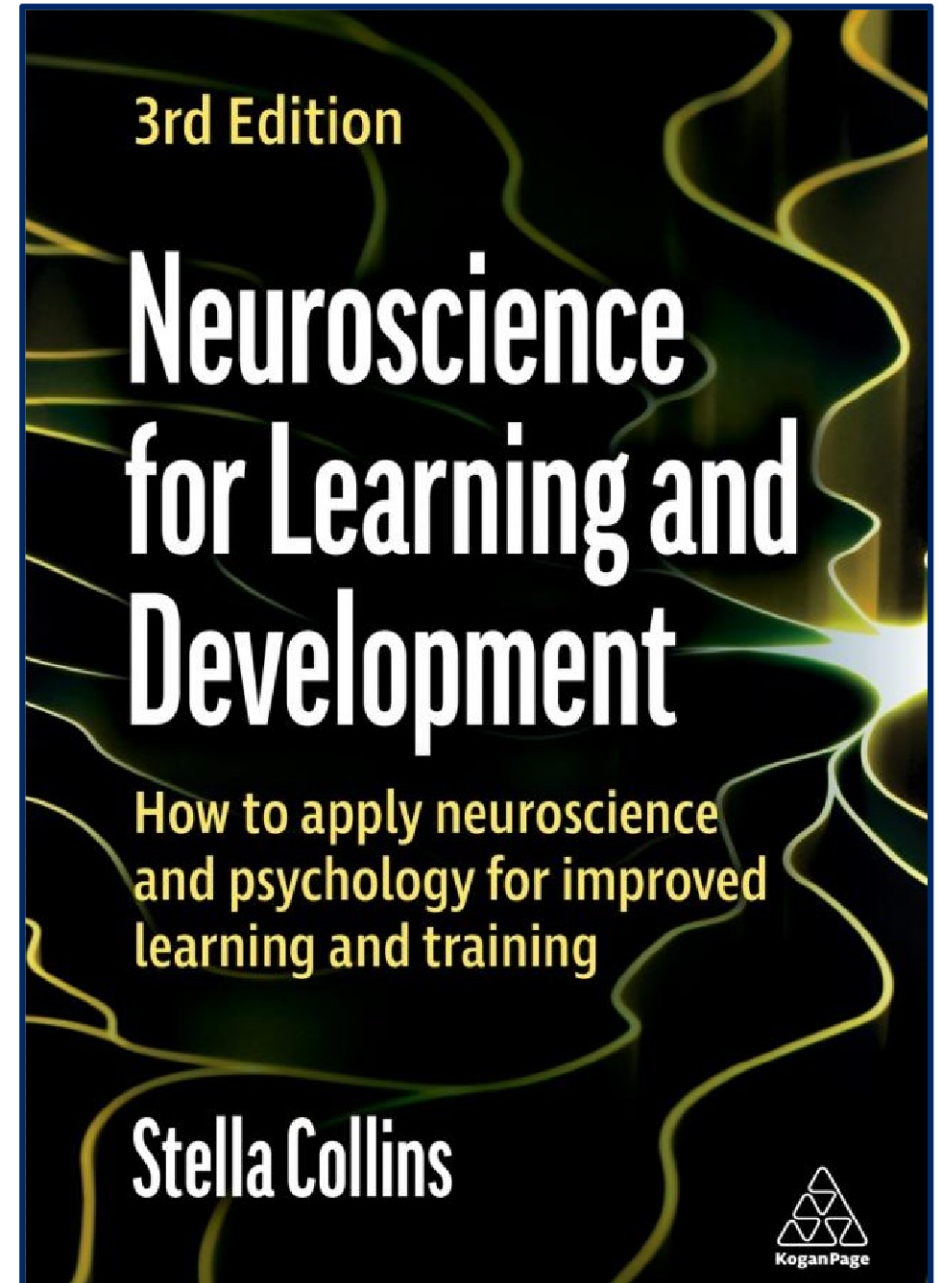
There is indeed evidence that fear appeals have been successful, but the arguments against their use, and the wide variety of experimental designs and evaluations, make it **very difficult to have confidence that they will prove efficacious in encouraging long-term secure behaviors.**

# Adult Learning & Training

## Chapter 5 – Motivation

How to motivate people to get their brains to learn:

- Curiosity
- Relaxation
- Persistence
- Goal orientation
- Creativity/playfulness



# Learning from Safety Culture

Safety, leadership and learning is based on **Human Organisational Performance (HOP)** and is a further development of our existing approach to safety.

## The HOP principles:

1. People make mistakes
2. Blame fixes nothing
3. Learning is the key to improvement
4. Context drives behaviour
5. How we respond matters

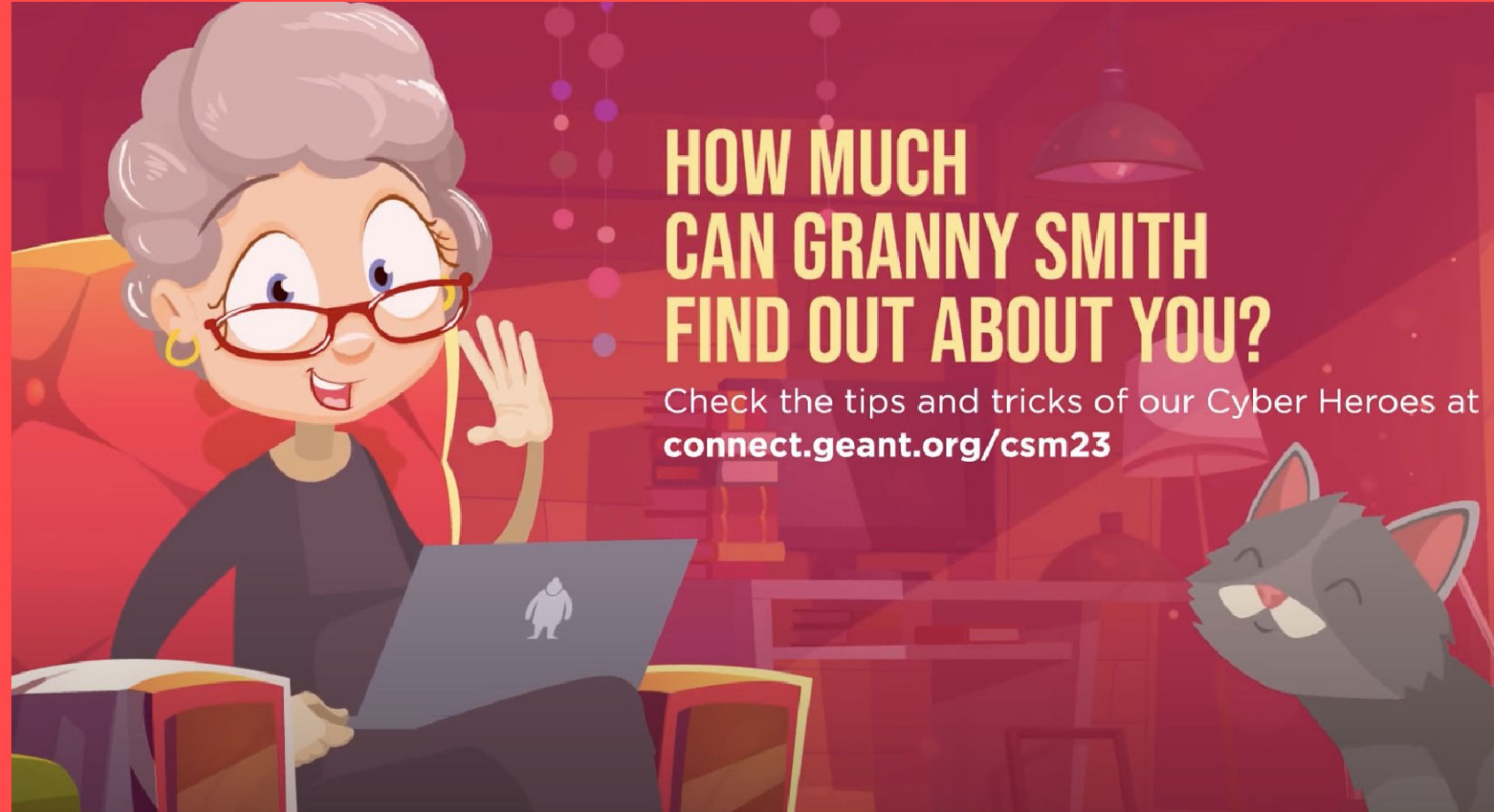


# Teaching the adversarial mindset

Cybercrime for  
Newbies with  
**#grannysmith85**

Mini video series on  
social engineering  
attack cycle

Switch\_

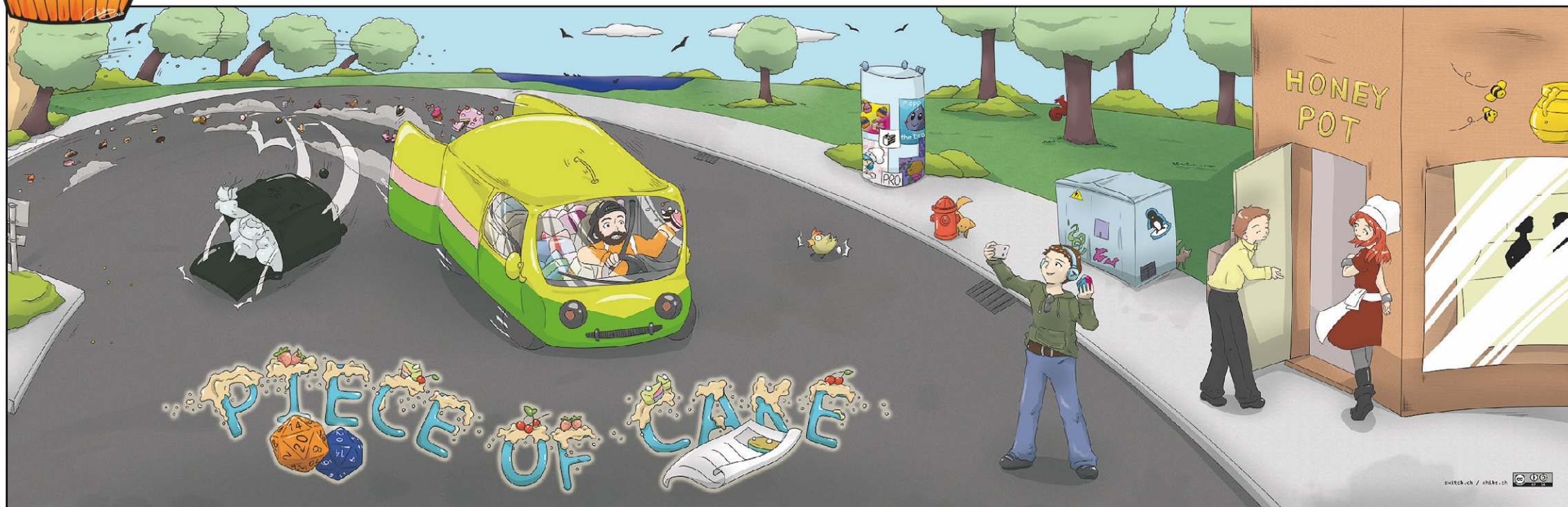


[https://www.youtube.com/watch?v=riwPE\\_qU7f0](https://www.youtube.com/watch?v=riwPE_qU7f0)

# Training the adversarial mindset





Piece of Cake – the role playing game





# Learning from Safety Culture

## Learning from safety science: A way forward for studying cybersecurity incidents in organizations

Nico Ebert<sup>a</sup>  , Thierry Schaltegger<sup>a</sup>, Benjamin Ambuehl<sup>a</sup>, Lorin Schöni<sup>b</sup>, Verena Zimmermann<sup>b</sup>, Melanie Knieps<sup>c</sup>

Show more 

+ Add to Mendeley  Share  Cite

<https://doi.org/10.1016/j.cose.2023.103435> 

[Get rights and content](#) 

Under a Creative Commons [license](#) 

 *open access*

Switch