June 2026

Deception By Design: Deepfakes And Malicious Insider Deviance In Cybersecurity.

Emmanuel Anti
University of Vaasa, emmanuel.anti@uwasa.fi

Duong Dang
University of Vaasa, duong.dang@uwasa.fi

Quang Bui
Rochester Institute of Technology, quang.bui@uwasa.fi

DECEPTION BY DESIGN: DEEPFAKES AND MALICIOUS INSIDER DEVIANCE IN CYBERSECURITY.

Short Paper

Emmanuel Anti, University of Vaasa, Finland, emmanuel.anti@uwasa.fi

Duong Dang, University of Vaasa, Finland, duong.dang@uwasa.fi

Quang "Neo" Bui, Rochester Institute of Technology, USA, qnbui@saunders.rit.edu

Abstract

Deepfake technology poses emerging risks for organizations by enabling the manipulation of audio, video, and images in ways that insiders can exploit to commit fraud, impersonate colleagues, or sabotage operations. This study extends Fraud Triangle Theory (FTT) to examine how deepfakes influence insider deviant behavior by reshaping perceptions of pressure, opportunity, and rationalization. This study will use quantitative methods to survey professionals across Europe, America, and Asia (n=250), testing a model of deepfake-enabled insider deviance while examining context-specific threats, motivations, and ethical rationalizations. Structural equation modeling will be used to validate and interpret findings. This study contributes to the IS literature by integrating emerging technologies into fraud theory, highlighting the misuse of deepfakes as a critical internal threat, and offering practical guidance for security governance, policy development, and AI-based detection strategies.

Keywords: Deepfakes, Malicious Insider Deviance, Fraud Triangle Theory, Cybersecurity

1 Introduction

As organizations increasingly rely on digital infrastructure, the cybersecurity threat landscape continues to evolve. Insider deviance, defined as unauthorized, unethical, or harmful behavior carried out by individuals with legitimate organizational access, remains one of the most difficult threats to detect and manage. Such actions exploit institutional trust, bypass traditional security controls, and can cause significant financial, operational, and reputational damage (Hunker & Probst, 2011; Liu et al., 2018).

Emerging technologies such as deepfakes add a new layer of complexity to this challenge. Deepfake technology uses machine learning techniques, particularly Generative Adversarial Networks, to generate highly realistic synthetic audio, video, or image content (Albahar & Almalki, 2019; Vyas et al., 2024). Once limited to specialized domains such as film production and national security (Temir, 2020), deepfake tools are now widely accessible through user-friendly software and mobile applications (Mahmud & Sharmin, 2021). Their capabilities, especially anonymity, plausibility, and deniability, make them particularly concerning in organizational settings because they enable deceptive actions while hiding attribution.

Unlike traditional cyberattacks that primarily exploit technical vulnerabilities, deepfakes manipulate human perception and trust. In a widely reported case, attackers used a voice-cloned deepfake of a chief executive officer to deceive an employee into transferring €220,000 to fraudulent accounts (Pedersen et al., 2025). Within organizations, insiders could similarly use deepfakes to impersonate colleagues, fabricate approvals, bypass biometric authentication mechanisms, or retaliate against perceived workplace injustices (Rini & Cohen, 2022). These developments raise important questions about how technological capabilities interact with behavioral motivations to produce malicious insider deviance.

Despite these risks, information systems security research has only begun to examine how generative AI technologies influence insider behavior. Existing work has focused more heavily on technical threats and external attackers, with less attention to the psychological and situational conditions through which insiders may misuse emerging technologies. To address this gap, we draw on Fraud Triangle Theory (FTT) (Albrecht et al., 1984, 2008), which explains deviant behavior through pressure, opportunity, and rationalization.

This study extends FTT by conceptualizing deepfake capabilities as conditions that strengthen the translation of these three drivers into malicious insider deviance. Specifically, deepfake-enabled anonymity (Chesney & Citron, 2019; Plohl et al., 2025; Hite et al., 2014) can reduce perceived attribution risk, plausibility (Chesney & Citron, 2019; Plohl et al., 2025) can increase the credibility of deception, and deniability (Chesney & Citron, 2019; Plohl et al., 2025) can weaken ethical accountability. These capabilities are therefore modelled as moderating conditions rather than primary antecedents, because they do not create pressure, opportunity, or rationalization themselves but alter the extent to which these drivers lead to deviant action.

Based on this perspective, the study addresses the following research questions:

RQ1: How do the elements of Fraud Triangle Theory, namely pressure, opportunity, and rationalization, influence malicious insider deviant behavior?

RQ2: How do deepfake capabilities, such as anonymity, plausibility, and deniability, condition the relationship between Fraud Triangle drivers and malicious insider deviance?

To answer these questions, we develop a conceptual model that integrates deepfake capabilities into the Fraud Triangle framework and propose a set of testable hypotheses. In doing so, the study contributes to information systems security research by showing that deepfakes are not merely new tools of deception but socio-technical conditions that can reshape how insider motivations translate into harmful behavior.

2 Theoretical Background

Malicious Insider Deviance and Deepfakes

Insider threats, whether malicious or non-malicious, remain a critical concern for organizations, particularly as emerging technologies expand the scale and sophistication of insider capabilities (Anti & Vartiainen, 2024). Individuals with privileged access, such as employees or contractors, can exploit internal systems to compromise sensitive data, sabotage operations, or manipulate organizational processes (Hunker & Probst, 2011; Liu et al., 2018). Such behaviors are described as insider deviant behaviors, defined as intentional employee actions that violate organizational norms and harm the organization or its stakeholders (Anti & Vartiainen, 2024). Recent industry reports indicate a rise in insider incidents from 66% in 2019 to 76% in 2024, underscoring the growing difficulty of detecting malicious insiders compared to external attackers (Cybersecurity Insiders, 2024). Insider deviance is especially difficult to manage because it exploits legitimate access, organizational trust, and knowledge of internal processes, driven by underlying psychological or organizational factors (Anti & Vartiainen, 2024).

Among emerging technologies, deepfakes represent a distinct and evolving threat vector. Deepfakes are synthetic media generated through artificial intelligence that can realistically manipulate audio, video, or image content to impersonate individuals or fabricate events (Chesney & Citron, 2019). Unlike traditional forms of insider deviance that rely on direct system or data manipulation, deepfakes enable deception by targeting human perception and trust. This makes them especially potent for impersonation, falsification of evidence, and manipulation of colleagues within organizational settings. For example, deepfake pornography has been used to harass or discredit employees in professional environments (Gieseke, 2020; Maddocks, 2020), while deepfake voice impersonation has enabled fraudulent financial transfers by mimicking executives (Juefei Xu et al., 2022).

Deepfakes are particularly significant because they introduce capabilities that conventional insider threat models do not fully capture. In this study, we focus on three such capabilities: anonymity, plausibility, and deniability. Anonymity reduces perceived attribution risk by making it harder to trace deceptive actions to the perpetrator (Chesney & Citron, 2019; Plohl et al., 2025; Hite et al., 2014). Plausibility increases the credibility of manipulated content by producing highly realistic synthetic media (Chesney & Citron, 2019; Plohl et al., 2025). Deniability allows actors to distance themselves from harmful actions by attributing suspicious content to technical artifacts, manipulation, or external interference (Chesney & Citron, 2019; Plohl et al., 2025). These capabilities together make deepfakes especially suitable for covert, persuasive, and difficult-to-attribute forms of insider deviance.

Existing insider deviance research has often relied on rational choice and deterrence perspectives that emphasize cost-benefit calculations or sanctions (Anti & Vartiainen, 2024). While valuable, these perspectives are less well-suited to explaining behavior shaped by psychological justification and technology-enabled deception. Deepfakes blur the boundary between digital manipulation and social influence by allowing insiders to fabricate convincing identities, interactions, and evidence. Understanding deepfake-enabled insider deviance, therefore, requires a framework that captures both the motivational pressures and the contextual conditions that drive misuse. To address this need, we draw on Fraud Triangle Theory as a lens for examining how technological capabilities interact with human motivations to produce insider deviant behavior.

2.1 Fraud Triangle Theory (FTT)

Fraud Triangle Theory (FTT), developed by Cressey (1950, 1953), explains fraud through three contextual conditions: pressure, opportunity, and rationalization. Over time, FTT has been expanded to incorporate the psychological and sociological dimensions of deviant behavior (Dorminey et al., 2012). Albrecht et al. (1984, 2008) further showed that fraud arises when individuals experience financial or non-financial pressures, perceive an opportunity due to weak oversight or control failures, and rationalize their actions as acceptable.

In FTT, pressure refers to personal, job-related, or external stressors such as financial hardship, career insecurity, or perceived injustice (Boal & Cummings, 1981; Lazarus & Folkman, 1984). Cultural or political conflict may further intensify pressure, leading to deviance motivated by survival, revenge, or gain (Chirasha & Mahapa, 2012). Opportunity reflects the perceived ability to commit fraud without detection and is shaped by access to systems, weak internal controls, and organizational complexity (Davis, 1989; Eisenhardt, 1985). Rationalization refers to the cognitive framing of unethical behavior as acceptable or justified, often influenced by moral reasoning and ethical climate (Siponen et al., 2020).

Recent work has begun extending fraud theory to AI-enabled deception. Zweers et al. (2025), for example, propose the AI-Fraud Diamond to explain new forms of deception and the auditing challenges they pose. However, such work focuses more on algorithmic deception than on the behavioral dynamics of insiders who intentionally misuse AI tools within organizations. FTT, therefore, remains particularly suitable here because it captures the motivational, situational, and justificatory conditions underlying insider deviance.

This study argues that FTT is especially well-suited to explaining deepfake-enabled insider deviance. Unlike traditional fraud that often involves direct manipulation of records or systems, deepfakes enable deception through fabricated identities, manipulated media, and persuasive synthetic content. As a result, they can reshape how insiders perceive risk, opportunity, and responsibility. Deepfakes may intensify pressure by enabling low-risk retaliation, sabotage, or manipulation in high-stress environments. They may expand opportunities by lowering the barriers to creating realistic false content and exploiting organizational trust. They may also strengthen rationalization by allowing perpetrators to frame their actions as harmless, deniable, or necessary, particularly when oversight is weak.

FTT has been used to examine how employees justify computer fraud and related unethical technology use (Jing, 2022; Owusu et al., 2022; Rahman & Jie, 2024). Building on this foundation, this study extends FTT by conceptualizing deepfake capabilities as a novel enabler of insider deviance. Specifically, deepfake-related anonymity, plausibility, and deniability are argued to condition how pressure, opportunity, and rationalization translate into malicious insider behavior. Table 1 outlines their distinct characteristics and shows how they may influence insider deviance within the Fraud Triangle framework. Although these characteristics may influence multiple elements of the Fraud Triangle, this study focuses on the pathway where each is expected to have the strongest effect. Next, we present the research model and hypotheses.

Deepfake Characteristic	Description	Implication for Insider Deviance	FTT Dimension
Anonymity (Chesney & Citron, 2019; Plohl et al., 2025; Hite et al., 2014).	Deepfakes can mask authorship or origin, reducing traceability	Lowers perceived attribution risk and may embolden insiders to act on existing motives under stress, grievances, or for retaliation.	Pressure
Plausibility (Chesney & Citron, 2019; Plohl et al., 2025)	Deepfakes create highly realistic synthetic media that can convincingly imitate people or events	Increases the credibility and success of deceptive acts such as impersonation or falsified approvals	Opportunity
Deniability (Chesney & Citron, 2019; Plohl et al., 2025)	Deepfakes create ambiguity about authenticity and responsibility	Makes it easier for perpetrators to distance themselves from harmful actions and justify misconduct, reducing accountability.	Rationalization

Table 1: Deepfake Characteristics and Their Implications for Insider Deviance

Hypothesis Development

Fraud Triangle Theory serves as the overarching framework for this study by explaining deviant behavior through the interaction of pressure, opportunity, and rationalization (Albrecht et al., 1984, 2008). To capture deepfake-enabled insider deviance more precisely, each FTT dimension is

operationalized through specific constructs reflecting organizational and technological conditions. Figure 1 presents the proposed research model.

Pressure

Pressure reflects internal or external stressors that motivate unethical behavior. In this study, pressure is operationalized through professional gains and personal grievances. Professional gains refer to motivations to obtain recognition, career advancement, or performance-related benefits through unethical means (Baskerville et al., 2014). Personal grievances refer to feelings of unfair treatment that motivate retaliatory deviant behavior (Connolly et al., 2017; Hina et al., 2019). In deepfake-enabled contexts, employees may use manipulated media to fabricate achievements, damage rivals, or retaliate against perceived mistreatment. Accordingly, we propose:

H1a: *Professional gains positively influence malicious insider deviant behavior.*

H1b: *Personal grievances positively influence malicious insider deviant behavior.*

Opportunity

Opportunity represents the perceived ability to act unethically without detection. We operationalize opportunity through data accessibility and tool accessibility. Data accessibility refers to access to privileged systems, information, or content that can be exploited for malicious purposes (Dhillon et al., 2020). Tool accessibility refers to the availability and ease of use of deepfake creation tools (Lin & Kunnathur, 2013). In combination, access to organizational resources and accessible generative tools can make deceptive actions more feasible. Here, opportunity captures access conditions within the organization, whereas deepfake capabilities capture the effectiveness of deception once such access exists. Thus, we propose:

H2a*: Data accessibility positively influences malicious insider deviant behavior.*

H2b*: Accessibility of deepfake tools positively influences malicious insider deviant behavior.*

Rationalization

Rationalization involves cognitive strategies that neutralize guilt and justify deviance. Building on prior literature (Anti & Vartiainen, 2024), we operationalize rationalization through moral disengagement and minimizing perceived harm. Moral disengagement refers to cognitive distancing from ethical standards that justifies misconduct (Chen et al., 2019). Minimizing perceived harm refers to the perception that deepfake misuse causes little or no real damage (Siponen & Vance, 2010). Because deepfakes often involve manipulation at a distance and can mask direct responsibility, insiders may frame their actions as harmless, necessary, or victimless, thereby reducing guilt and enabling deviance (Chen et al., 2019; Siponen & Vance, 2010). Hence, we propose:

H3a: *Moral disengagement positively influences malicious insider deviant behavior.*

H3b: *Minimizing perceived harm positively influences malicious insider deviant behavior.*

Moderating Role of Deepfake Capabilities

Deepfake-related capabilities do not directly create pressure, opportunity, or rationalization; instead, they shape how strongly these drivers are translated into malicious insider deviant behavior. Anonymity reduces traceability and perceived detection risk, making individuals more willing to act on existing motives for professional gain or personal grievance (Hite et al., 2014). For example, an insider seeking promotion or retaliation may be more willing to send manipulated content if they believe the act cannot easily be traced back to them. Plausibility increases the credibility of manipulated content, making existing opportunities easier to exploit through more convincing deception (Chesney & Citron, 2019; Plohl et al., 2025). For example, a realistic synthetic voice or video can make a fraudulent request appear legitimate enough to bypass normal suspicion. Deniability reduces perceived accountability by making authorship easier to dispute, thereby supporting post hoc justification (Chesney & Citron, 2019; Plohl et al., 2025). For example, an insider may later claim that the manipulated message was fake, misunderstood, or not directly attributable to them. Accordingly, these capabilities are modelled as moderators, as they amplify rather than replace the effects of the core fraud drivers. Accordingly, we propose:

***H4a**: Deepfake-related anonymity strengthens the relationship between pressure and malicious insider deviant behavior.*

***H4b**: Deepfake-related plausibility strengthens the relationship between opportunity and malicious insider deviant behavior.*

***H4c**: Deepfake-related deniability strengthens the relationship between rationalization and malicious insider deviant behavior.*

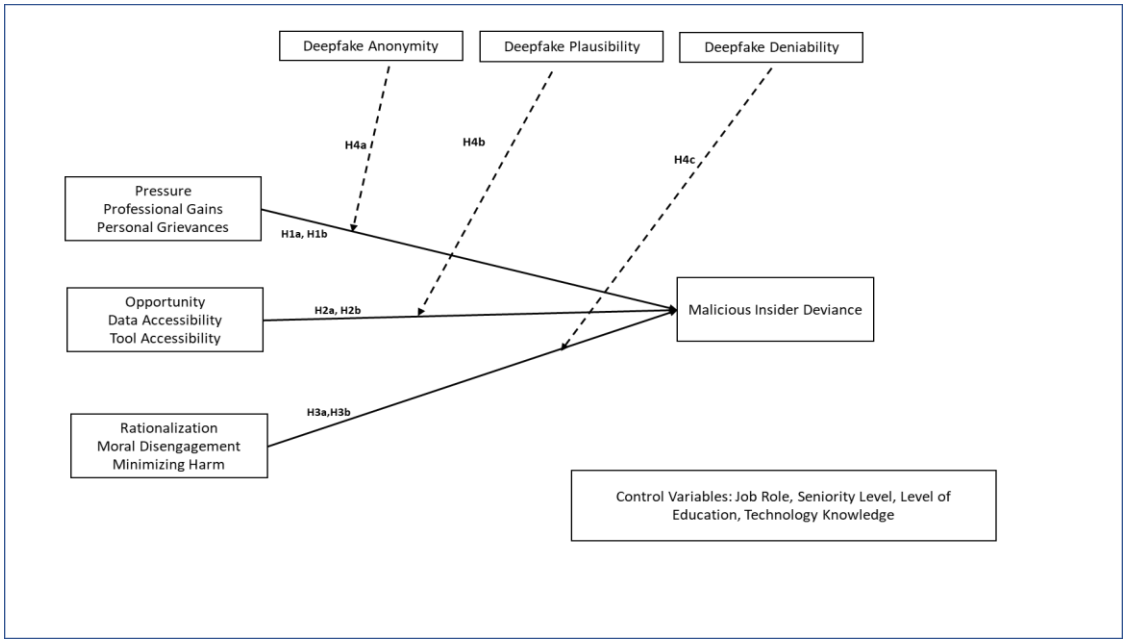

Figure 1: Proposed model of deepfake-enabled malicious insider deviance

3 Methodology

This study adopts a quantitative survey design to examine deepfake-enabled insider deviant behavior through the lens of Fraud Triangle Theory (FTT). A survey is appropriate for testing the proposed relationships and generating initial empirical evidence on an emerging phenomenon (Ivankova & Creswell, 2009; Yilmaz, 2013). The instrument was developed by adapting validated measures from prior research and refining them for the deepfake context (Roopa & Rani, 2012). A pilot study with 20 participants, together with feedback from two academic experts and a doctoral seminar, was used to assess item clarity and construct relevance prior to full data collection.

All variables will be measured using multi-item seven-point Likert scales. The target sample comprises 250 participants from Asia, Europe, and North America, recruited through stratified random sampling to ensure representation across industries and professional roles, and regions (Singh & Masuku, 2014). Data will be analyzed using structural equation modeling, with confirmatory factor analysis used to assess construct validity. Reliability and discriminant validity will be evaluated following Sarstedt et al. (2021) and Fornell and Larcker (1981). Control variables include job role, seniority, level of education, and technology knowledge, while common method bias will be mitigated using procedural and statistical remedies (Podsakoff et al., 2003; Henseler, 2007). As the empirical study is still in progress, the survey is intended to provide first-stage evidence on this phenomenon and to inform future, more behaviorally grounded designs.

4 Expected Contributions and Conclusion

This study is expected to extend the Fraud Triangle Theory to the context of deepfake-enabled insider deviance. Rather than treating deepfakes as merely new tools of deception, it conceptualizes them as socio-technical capabilities that condition how pressure, opportunity, and rationalization translate into

malicious insider behavior. In doing so, it brings emerging generative AI risks into an established behavioral framework for insider deviance.

The study is expected to offer practical insight into how organizations can respond to deepfake-enabled insider threats. By showing how deepfake capabilities may interact with insider motivations, it may inform governance measures, monitoring approaches, and employee awareness efforts related to identity verification, synthetic media risks, and internal security practices.

Work Remaining

Quantitative data collection is scheduled to begin in April 2026, following incorporation of reviewer feedback and refinement of the instrument. Data collection is expected to be completed by July 2026, with analysis to follow by October 2026 and the final paper by December 2026. Future research may build on these survey findings through scenario-based or experimental designs.

References

Albahar, M., & Almalki, J. (2019). Deepfakes: Threats and countermeasures systematic review. Journal of Theoretical and Applied Information Technology, 97(22), 3242–3250.

Albrecht, W. S., Albrecht, C., & Albrecht,C. C. (2008). Current trends in fraud and its detection. Information Security Journal: A Global Perspective, 17(1), 2–12.

Albrecht, W. S., Howe, K. R., & Romney, M. B. (1984). Deterring fraud: The internal auditor's perspective.

Anti, E., & Vartiainen, T. (2024). Explanations of Insider Deviant Behavior in Information Security: A Systematic Literature Review. Communications of the Association for Information Systems, 55(1), 4.

Baskerville, R., Park, E. H., & Kim, J. (2014). An emote opportunity model of computer abuse. Information Technology & People, 27(2), 155–181. doi:10.1108/ITP-11-2011-0068.

Boal, K. B., & Cummings, L. (1981). Cognitive evaluation theory: An experimental test of processes and outcomes. Organizational Behavior and Human Performance, 28(3), 289–310.

Box, D., & Pottas, D. (2014). A model for information security compliant behaviour in the healthcare context. Procedia Technology, 16, 1462–1470.

Chen, H., Chau, P. Y., & Li, W. (2019). The effects of moral disengagement and organizational ethical climate on insiders' information security policy violation behavior. Information Technology & People, 32(4), 973–992.doi:10.1108/ITP-12-2017-0421

Chesney, B., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. Calif. L. Rev., 107, 1753.

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. Foreign Aff., 98, 147.

Chirasha, V., & Mahapa, M. (2012). An analysis of the causes and impact of deviant behaviour in the workplace. The Case of Secretaries in State Universities. Journal of Emerging Trends in Economics and Management Sciences, 3(5), 415–421.

Connolly, Y., L., L., M., G., J., & Tygar, D. J. (2017). Organisational culture, procedural countermeasures, and employee security behaviour: A qualitative study. Information & Computer Security, 25(2), 118–136.

Cressey, D. R. (1950). The Criminal Violation of Financial Trust. American Sociological Review, 15(6), 738-743.

Cressey, D. R. (1953). Other people's money; a study of the social psychology of embezzlement.

Cybersecurity Insiders. (2024). Insider Threat Report Trends, Challenges, and Solution (p. 26). www.securonix.com

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The Development and Validation of a New Machiavellianism Scale Journal of Management, 35(2), 219–257.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 319–340.

Dhillon, G., Abdul Talib, Y. Y., & Picoto, W. N. (2020). The mediating role of psychological empowerment in information security compliance intentions. Journal of the Association for Information Systems, 21(1), 5.

Dorminey, J., Fleming, A. S., Kranacher, M.-J., & Riley Jr, R. A. (2012). The Evolution of Fraud Theory. Issues in Accounting Education, 27(2), 555.

Eisenhardt, K. M. (1985). Control: Organizational and economic approaches. Management Science, 31(2), 134–149.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50.

Gieseke, A. P. (2020). "The New Weapon of Choice": Law's Current Inability to Properly Address Deepfake Pornography. Vand. L. Rev., 73, 1479.

Henseler, J. (2007). A new and simple approach to multi-group analysis in partial least squares path modeling. 5th International Symposium on PLS and Related Methods, PLS 2007: Causalities Explored by Inditect Observation, 104–107.

Hina, S., Selvam, D. D. D. P., & Lowry, P. B. (2019). Institutional governance and protection motivation: Theoretical insights into shaping employees' security compliance behavior in higher education institutions in the developing world. Computers & Security, 87, 101594.

Hite, D. M., Voelker, T., & Robertson, A. (2014). Measuring perceived anonymity: The development of a context-independent instrument. Journal of Methods and Measurement in the Social Sciences, 5(1), 22–39.

Hunker, J., & Probst, C. W. (2011). Insiders and Insider Threats-An Overview of Definitions and Mitigation Techniques. J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl., 2(1), 4–27.

Ivankova, N. V., & Creswell, J. W. (2009). Mixed methods. Qualitative Research in Applied Linguistics: A Practical Introduction, 23, 135–161.

Jiang, R. (2022). Exploring Employees' Computer Fraud Behaviors using the Fraud Triangle Theory. Pacific Asia Journal of the Association for Information Systems, 14(4), 4.

Juefei-Xu, F., Wang, R., Huang, Y., Guo, Q., Ma, L., & Liu, Y. (2022). Countering Malicious Deepfakes: Survey, Battleground, and Horizon. International Journal of Computer Vision, 130(7), 1678–1734.

Lazarus, R. S., & Folkman, S. (1984). Stress, appraisal, and coping. Springer publishing company.

Liu, L., De Vel, O., Han, Q.-L., Zhang, J., & Xiang, Y. (2018). Detecting and preventing cyber insider threats: A survey. IEEE Communications Surveys & Tutorials, 20(2), 1397–1417.

Lin, C., & Kunnathur, A. S. (2013). Toward Developing a Theory of End User Information Security Competence.

Maddocks, S. (2020). 'A Deepfake Porn Plot Intended to Silence Me': Exploring continuities between pornographic and 'political'deep fakes. Porn Studies, 7(4), 415–423.

Mahmud, B. U., & Sharmin, A. (2021). Deep insights of deepfake technology: A review. arXiv Preprint arXiv:2105.00192.

Nam, S.-J. (2023). Deviant behavior in cyberspace and emotional states. Current Psychology, 42(13), 10751–10760.

Owusu, G. M. Y., Koomson, T. A. A., Alipoe, S. A., & Kani, Y. A. (2022). Examining the predictors of fraud in state-owned enterprises: An application of the fraud triangle theory. Journal of Money Laundering Control, 25(2), 427–444.

Pedersen, K. T., Pepke, L., Stærmose, T., Papaioannou, M., Choudhary, G., & Dragoni, N. (2025). Deepfake-Driven Social Engineering: Threats, Detection Techniques, and Defensive Strategies in Corporate Environments. Journal of Cybersecurity and Privacy, 5(2). https://doi.org/10.3390/jcp5020018

Plohl, N., Mlakar, I., Aquilino, L., Bisconti, P., & Smrke, U. (2025). Development and Validation of the Perceived Deepfake Trustworthiness Questionnaire (PDTQ) in Three Languages. International Journal of Human–Computer Interaction, 41(11), 6786–6803.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879.

Rahman, M. J., & Jie, X. (2024). Fraud detection using fraud triangle theory: Evidence from China. Journal of Financial Crime, 31(1), 101–118.

Rini, R., & Cohen, L. (2022). Deepfakes, deep harms. J. Ethics & Soc. Phil., 22, 143.

Roopa, S., & Rani, M. S. (2012). Questionnaire designing for a survey. Journal of Indian Orthodontic Society, 46(4_suppl1), 273–277.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In Handbook of market research (pp. 587–632). Springer.

Singh, A. S., & Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: An overview. International Journal of Economics, Commerce and Management, 2(11), 1–22.

Siponen, M., Puhakainen, P., & Vance, A. (2020). Can individuals' neutralization techniques be overcome? A field experiment on password policy. Computers & Security, 88, 101617. https://doi.org/10.1016/j.cose.2019.101617

Siponen, M., & Vance, A. (2010). Neutralization: New insights into the problem of employee information systems security policy violations. MIS Quarterly, 487–502.doi:10.2307/25750688

Temir, E. (2020). Deepfake: New era in the age of disinformation & end of reliable journalism. Selçuk İletişim, 13(2), 1009–1024.

Vyas, K., Pareek, P., Jayaswal, R., & Patil, S. (2024). Analysing the landscape of Deep Fake Detection: A Survey. International Journal of Intelligent Systems and Applications in Engineering, 12(11s), 40–55.

Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. European Journal of Education, 48(2), 311–325.

Zweers, B., Dey, D., & Bhaumik, D. (2025). The AI-Fraud Diamond: A Novel Lens for Auditing Algorithmic Deception. arXiv Preprint arXiv:2508.13984.