AB Journals

British Journal of Computer, Networking and Information Technology

ISSN 2689-5315 • Volume 9, Issue 3, 2026 • pp. 60–83

RESEARCH ARTICLE Received: 6 Oct 2026, Accepted: 11 Nov 2026, Published: 19 Nov 2026

Implementation, Validation, and Benchmarking of a Composite Framework for Evaluating Firewall and Network Access Control Effectiveness

Dr. Funmilayo Abibat Sanusi[1*], Anita O. Onyeacholem[2], Victor C. Ezeokechukwu[2], Grace Chinenye Okorie[1], Dr. Mensah Yaw[1]

[1]Department of Software Engineering, School of Computing, Babcock University, Ilishan-Remo, Ogun State, Nigeria.
[2]Department of Computer Science, School of Computing, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

[*]*Corresponding Author's Email: sanusifu@babcock.edu.ng*

Abstract

Nigerian banks have invested heavily in firewalls and Network Access Control (NAC). However, few instruments test whether these controls work, and fewer are validated or benchmarked against recognised standards. This paper therefore specifies how a composite framework that scores firewall and NAC effectiveness will be implemented, validated, and benchmarked. The study adopts a pragmatic, convergent mixed-methods design within a design science approach. Three instruments serve as base models: NIST SP 800-41 Revision 1, PCI DSS v4.0 Requirements 1 and 7, and a published banking risk framework. Implementation follows a five-input pipeline, and where self-reports and documents disagree, the lower value is used. Validation combines an expert panel, the Delphi technique, scenario-based testing, and benchmark comparison. The design yields three falsifiable benchmarking tests: coverage completeness, measurement extension, and contextual adaptation. It also traces every firewall and NAC dimension to a documented requirement, and it turns the integration claim into a testable interaction term. The study concludes that a security evaluation framework can be designed to fail visibly. As a result, its eventual findings will be meaningful whether they are positive or negative.

Keywords: Benchmarking, Firewall Evaluation, Framework Validation, Network Access Control, Security Metrics

How to Cite: Sanusi, F. A., Onyeacholem, A. O., Ezeokechukwu, V. C., Okorie, G. C., & Yaw, M. (2026). Implementation, Validation, and Benchmarking of a Composite Framework for Evaluating Firewall and Network Access Control Effectiveness. *British Journal of Computer, Networking and Information Technology*, 9(3), 60–83. https://doi.org/10.52589/BJCNIT-DRIIKZ

© 2026 by the authors. Published by AB Journals. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) license.

INTRODUCTION

Banking has shifted decisively from the branch to digital channels, and Nigeria is part of that shift. Hassan et al. (2024) note that banking digitisation brings cybersecurity challenges ranging from data breaches and ransomware attacks to sophisticated financial fraud, and that Nigerian banks face distinctive challenges of their own. The pattern is international. Financial technology adoption has been shown to influence the cybersecurity performance of listed financial institutions (Gyamfi et al., 2026). Similarly, the digital transformation of banking raises new data privacy and cybersecurity challenges (Wang et al., 2024). Each new digital channel therefore widens the surface that an attacker can reach.

The consequences of failure are costly. The global average cost of a data breach reached USD 4.88 million in 2024, and the financial sector averaged USD 6.08 million (IBM, 2024). The same report found that financial services organisations took an average of 168 days to identify a breach and a further 51 days to contain it. Research also links cyber-attacks and cybersecurity risk to banks' intermediation and competition (Boungou, 2023; Fuszder et al., 2026). In response, the Central Bank of Nigeria (CBN, 2024) issued a risk-based cybersecurity framework. The framework sets minimum cybersecurity requirements for deposit money banks and payment service banks. It also requires documented self-assessment, which the CBN reviews through risk-based supervision.

Firewalls and Network Access Control (NAC) sit at the base of this defence. A firewall filters traffic at network boundaries according to rules that the institution defines (Scarfone & Hoffman, 2009). NAC, in turn, decides which users and devices may join the network and what they may reach once admitted. The Payment Card Industry Data Security Standard (PCI DSS) treats the two functions as a matched pair: Requirement 1 governs network security controls, while Requirement 7 restricts access according to business need (PCI Security Standards Council [PCI SSC], 2022). Owning these controls, however, is not the same as running them well. Firewall rule bases, for instance, accumulate anomalies through which one rule can silently override another and defeat the intended policy (Al-Shaer & Hamed, 2004).

Proposing a security evaluation framework is the easy part. Showing that it works is harder, and the literature is noticeably thinner there. Reviews of cybersecurity assessment methods repeatedly find instruments that are argued from principle rather than tested against anything, and frameworks validated only by the authors who built them (e.g., Leszczyna, 2021). A framework that has never been benchmarked against an established standard cannot claim to improve on it. Likewise, one that has never been applied outside a laboratory has not shown that it survives contact with an operating network.

Recent studies agree on three points. First, cyber risk in financial institutions is increasingly treated as a governance and quantification problem, addressed through integrated risk management frameworks and financial models of cyber risk (Azura et al., 2025; Osman, 2026; Pollmeier et al., 2023). Second, the quantification of controls is feasible. Configuration settings, access-control policies, and individual access decisions have all been expressed as computed scores (Belhaouane et al., 2015; Bradatsch et al., 2023; Scarfone & Mell, 2010). Third, perimeter controls alone are no longer considered sufficient. Zero Trust research stresses continuous authentication and granular access control (Gambo & Almulhem, 2026), and firewall research concedes that no single inspection method reliably identifies the endpoint behind the traffic (Heino et al., 2022).

The studies diverge, however, on where assurance should come from. One stream builds institution-wide models that quantify exposure across the banking technology stack (Azura et al., 2025; Pollmeier et al., 2023). A second stream redesigns architecture around Zero Trust, and it has validated its proposals through demonstration applications, prototypes, or algorithm designs rather than in operating banks (Chaudhry & Hydros, 2023; Daah et al., 2023, 2024). A third stream scores controls, but each study scores one element in isolation, whether a configuration setting, an access-control policy, or an individual access decision (Belhaouane et al., 2015; Bradatsch et al., 2023; Scarfone & Mell, 2010). The streams also differ in their evidence. Model-building studies rarely test deployed controls, while studies of operating banks, such as surveys of Nigerian bank staff, measure cybersecurity through broad proxies rather than configuration (Akintoye et al., 2022). A recent bibliometric analysis confirms the pattern: the field is expanding rapidly but remains fragmented, and its technical research streams are weakly integrated with governance and regulation (Sopko & Šafár, 2026).

Three gaps follow from this synthesis. First, no existing instrument scores firewall and NAC effectiveness as a coordinated pair. Second, few proposed frameworks are benchmarked against recognised standards or validated beyond their designers. Third, little of this work is calibrated to the regulatory conditions of developing economies such as Nigeria, where a binding framework requires self-assessment but does not prescribe configuration-level measures for individual controls (CBN, 2024).

This paper addresses these gaps for a specific instrument. A companion paper specifies a composite framework that scores firewall and NAC effectiveness numerically, combining a threat-level effectiveness measure, a rule-base complexity index, and a six-dimension NAC strength score into a single institutional index (Onyeacholem & Sanusi, 2026). The present paper sets out how that framework will be implemented, how it will be validated, and how it will be benchmarked against the instruments it claims to extend.

The study is necessary for both practical and scholarly reasons. In practice, banks that must evidence their controls to a regulator need measures that show how well a control performs, not merely that it exists. A validated instrument gives that evidence a defensible basis. For scholarship, the study advances the field in two ways. It makes the defence-in-depth claim that firewall and NAC are complementary testable rather than assumed. It also offers a replicable protocol for validating security evaluation frameworks in operating banks.

Three commitments shape the design. The framework is positioned against existing work rather than presented as an original departure, because evaluation frameworks for network security controls are not new and the claim to novelty lies in a narrower place than the whole instrument. Benchmarking is specified as a set of tests with pass conditions rather than as a general comparison, so that failure is possible and visible. And the validation design is stated before execution, so that the criteria cannot drift to accommodate whatever the results turn out to be.

The remainder of the paper is organised as follows. Section 2 reviews the theoretical foundations, synthesises the empirical literature, presents the base models, and introduces the conceptual framework. Section 3 sets out the methodology, including the research design, sampling, implementation, benchmarking, validation, and analytical procedures. Section 4 reports the design outcomes, and Section 5 discusses them against recent literature. Section 6 draws implications for research and practice. Section 7 states the limitations, Section 8 proposes directions for future research, and Section 9 concludes.

LITERATURE REVIEW

This section first sets out the theories that underpin the framework. It then synthesises recent empirical work to locate the gaps the study addresses, describes the base models used for benchmarking, and presents the conceptual framework.

Theoretical Framework

Three theories inform the study. Defence-in-depth explains what is evaluated, risk management theory explains how effectiveness is judged, and design science explains how the evaluation instrument itself should be validated.

Defence-in-Depth Theory

Defence-in-depth holds that protection should be layered and that the layers should be independent, so that breaching one does not expose the assets behind it (Kuipers & Fabro, 2006). Applied to networks, the theory favours diverse controls: firewalls at the boundary, access control and segmentation inside the network, and protection at the host. Its benefit is resilience. Its risk is administrative, because layers managed in isolation can overlap, enforce policy inconsistently, and create a false sense of security. For this study the theory does specific work. It treats the firewall and NAC as complementary layers whose coordinated effect should exceed what either delivers alone. That proposition underlies both the decision to score the pair jointly and the interaction term specified in Section 3.6.1.

Risk Management Theory

Risk management theory holds that risk should be managed through a structured and continuous process. That process establishes the context, identifies and analyses risks, selects treatments, and monitors whatever risk remains (International Organization for Standardization [ISO], 2018). Its scientific foundation lies in the systematic characterisation of consequences and their uncertainty (Aven, 2016). Applied to cybersecurity, the theory reframes what a control is. Firewalls and NAC become risk treatments, and their adequacy is judged against the likelihood and impact of the threats they address. This framing licenses the study's central move: effectiveness is treated as measurable risk reduction against defined indicators, not as a yes-or-no question of whether a control is installed. The theory's weakness is quantification, since threat data are incomplete and some consequences resist valuation. The sensitivity analysis described in Section 3.5.3 responds directly to that weakness.

Design Science and Artefact Evaluation

Design science research treats the building and evaluation of artefacts as a legitimate route to knowledge, provided the artefact is evaluated rigorously against the problem it was built to solve (Hevner et al., 2004). Process models in this tradition place demonstration and evaluation as explicit stages after design and development (Peffers et al., 2007). The Framework for Evaluation in Design Science (FEDS) adds that evaluation should be planned deliberately with respect to why, when, how, and what is evaluated (Venable et al., 2016). This theory governs the present paper most directly. It explains why the framework is developed iteratively, why validation combines several methods, and why benchmarking against existing artefacts forms part of evaluation.

Read together, the three theories close one another's gaps. Defence-in-depth justifies layering but says little about measurement. Risk management theory justifies measurement but is silent

on how an instrument should be validated. Design science supplies that validation logic. Zero Trust, formalised in National Institute of Standards and Technology (NIST) Special Publication (SP) 800-207, provides the horizon against which current deployments can be judged. It requires access decisions to be evaluated dynamically, per session, from identity and device posture (Rose et al., 2020). Recent reviews show that Zero Trust adoption still faces practical barriers (Gambo & Almulhem, 2026). The evaluation of existing perimeter and admission controls therefore remains a continuing need.

Empirical Review and Research Gaps

Recent empirical work falls into three streams. This section synthesises what each stream has found and what it leaves unaddressed, rather than describing individual studies in detail.

Governance, Quantification, and Banking Outcomes

The first stream treats cyber risk in financial institutions as a governance and measurement problem. Integrated frameworks now extend risk governance across human and machine identities (Osman, 2026). Quantitative models express cyber risk in financial or risk terms for bank decision-makers (Azura et al., 2025; Pollmeier et al., 2023). Econometric work, meanwhile, links cyber-attacks and cybersecurity risk to banks' intermediation and competition (Boungou, 2023; Fuszder et al., 2026). These studies align in showing that cyber risk is consequential and measurable at the level of the institution. They diverge in the level of analysis they favour, from financial loss models to identity governance. None, however, measures how well the technical controls that banks operate actually perform.

Firewall and Access-Control Evaluation

The second stream addresses the controls themselves. Firewall research shows that configuration quality matters as much as capability. Rule bases accumulate anomalies such as shadowing and redundancy (Al-Shaer & Hamed, 2004), and no single inspection method reliably identifies the endpoint application behind network traffic (Heino et al., 2022). That limitation supports pairing the firewall with an admission control. Access-control research, for its part, shows that policies and access decisions can be quantified. Examples include complexity metrics for authorisation policies (Belhaouane et al., 2015) and trust scores computed from quantified trust attributes (Bradatsch et al., 2023). Zero Trust proposals for financial institutions integrate identity, device, and network security. However, they have been evaluated through a demonstration bank application, a prototype, or an algorithm design rather than in operating banks (Chaudhry & Hydros, 2023; Daah et al., 2023, 2024). This stream therefore aligns on the feasibility of measurement. It diverges on the unit of measurement, from a single rule set to a single access request, and no study scores the firewall and NAC as one deployment.

The Nigerian and Developing-Economy Context

The third stream concerns context. Reviews of Nigerian banking describe rapid digitisation alongside distinctive cybersecurity challenges and locally developed defensive strategies (Hassan et al., 2024). Survey evidence from 56 purposively selected staff of Nigerian deposit money banks found that cybersecurity, measured through risk management and bank monitoring, had a significant positive effect on financial innovation (Akintoye et al., 2022). At the regulatory level, the CBN (2024) framework makes cybersecurity programmes mandatory and requires documented self-assessment. A bibliometric review of the wider field adds that

technical research remains weakly integrated with governance and regulation (Sopko & Šafár, 2026). The Nigerian evidence thus aligns with international work on the importance of cyber risk, but it rarely reaches the level of the individual control.

Synthesis: Research Gaps and the Study's Response

Table 1 draws the three streams together. It identifies four gaps and shows how the present study responds to each. The integration gap concerns the absence of joint firewall and NAC evaluation. The validation gap concerns frameworks that are proposed but neither benchmarked nor tested in operating banks. The contextual gap concerns the lack of calibration to Nigerian regulation. Finally, the measurement gap concerns reliance on self-report, which cannot detect the distance between what an institution believes its controls do and what its configuration shows.

Table 1: **Research Gaps and the Study's Response**

Gap	What recent studies show	How this study responds
Integration	Studies quantify one element at a time: rule sets, authorisation policies, or access requests (Al-Shaer & Hamed, 2004; Belhaouane et al., 2015; Bradatsch et al., 2023)	Scores firewall and NAC as a coordinated pair and tests their interaction (Section 3.6.1)
Validation	Zero Trust frameworks for banks are evaluated through demonstration, prototype, or design only (Chaudhry & Hydros, 2023; Daah et al., 2023, 2024)	Three benchmarking tests with pass conditions (Section 3.4) and four complementary validation methods (Section 3.5)
Context	Nigerian studies use broad cybersecurity proxies (Akintoye et al., 2022); technical research is weakly linked to regulation (Sopko & Šafár, 2026)	Contextual adaptation test calibrated to the CBN (2024) requirements
Measurement	Self-report measures are susceptible to response bias (Podsakoff et al., 2003)	Evidence reconciliation: documentary evidence overrides a higher self-report (Section 3.3.3)

Note: *The table synthesises the gaps identified in Sections 2.2.1 to 2.2.3. CBN = Central Bank of Nigeria; NAC = network access control.*
Source: *Author-generated (2026) from the studies cited in the table.*

Base Models for Benchmarking

Evaluation frameworks for network security controls already exist, and a review of the available instruments establishes not an absence but a mismatch. The instruments rich in configuration detail treat firewall and NAC separately. Those that treat traffic and access as a pair confine themselves to a narrow scope. None is calibrated to the regulatory environment of a Nigerian deposit money bank. A framework built into that space must therefore show what it adds to what already exists, which requires a baseline. Three instruments serve as base models, chosen because between them they supply technical criteria, paired control logic, and sector relevance. Table 2 compares them.

Table 2: Base Models and Their Roles in Benchmarking

Base model	Nature	Coverage of firewall and NAC	Role in benchmarking
NIST SP 800-41 Revision 1 (Scarfone & Hoffman, 2009)	Technical guideline	Comprehensive firewall coverage: technologies, policy, rules, placement, and lifecycle; NAC not addressed	Supplies the firewall criteria, the reference configuration, and the firewall benchmark
PCI DSS v4.0, Requirements 1 and 7 (PCI SSC, 2022)	Prescriptive standard with defined testing procedures; externally assessed	Requirement 1 covers network security controls; Requirement 7 restricts access by business need	Supplies the paired traffic-and-access logic in a form suitable for scoring
Integrated banking risk framework (Azura et al., 2025)	Published research framework	Treats network controls as one component of an institution-wide risk model	Sector comparator and standard of evidence for applicability

Note: *NIST = National Institute of Standards and Technology; SP = Special Publication; PCI DSS = Payment Card Industry Data Security Standard; PCI SSC = PCI Security Standards Council.*
Source: *Author-generated (2026), synthesised from Scarfone and Hoffman (2009), PCI SSC (2022), and Azura et al. (2025).*

NIST SP 800-41 Revision 1 as the Firewall Base Model

Scarfone and Hoffman (2009) provide the most widely adopted technical treatment of firewall policy and configuration, covering firewall technologies, policy formulation, rule construction, topological placement, and lifecycle management. Its role here is threefold. It supplies the firewall dimensions of the operationalisation presented in Section 4.1, so that every questionnaire item traces to a documented requirement. It provides the reference configuration against which observed practice is compared, so that a deviation can be identified as departure from a recognised standard rather than from the researcher's preference. And it constitutes the benchmark against which the firewall component of the framework is judged.

PCI DSS v4.0 as the Access Control Base Model

PCI DSS v4.0 serves as the second base model, with Requirements 1 and 7 carrying that role (PCI SSC, 2022). Two properties recommend it. It is prescriptive and externally assessed, each requirement carrying defined testing procedures, which means its criteria are already expressed in a form suitable for scoring. More importantly, it is the only widely adopted standard that treats traffic filtering and access restriction as a matched pair, which is structurally identical to the firewall and NAC relationship the framework addresses. Requirement 1 governs what may cross the boundary; Requirement 7 governs who may act once inside.

The Sector Base Model

Azura et al. (2025) developed an integrated cyber security risk management framework for

online banking systems, combining threat modelling with quantitative risk assessment across the banking technology stack. It is the closest published analogue to the framework proposed here, and it serves as the third base model for three reasons. It is set in banking rather than in a generic enterprise, so its control assumptions match the operating conditions examined. It expresses risk quantitatively rather than descriptively, which makes direct comparison possible. And its applicability was evaluated against the conditions of real-world online banking, including integration with existing technologies and regulatory requirements, rather than argued from principle alone. That evaluation supplies a standard of evidence the present framework is expected to meet.

Two differences are carried forward deliberately. The Azura framework treats network controls as one component within an institution-wide risk model, whereas the present framework narrows the focus to the firewall and NAC pair and examines that pair in greater depth. It also assumes a regulatory context other than Nigeria's. Benchmarking against it therefore tests whether narrowing scope and recalibrating to the environment set by the CBN and the Nigeria Data Protection Commission (NDPC) yields evaluative precision that a broader framework does not provide.

Conceptual Framework

Figure 1 presents the conceptual framework. It links the three theories to the constructs that the framework measures and to the procedures that test it. Firewall effectiveness and NAC effectiveness are the independent constructs. Drawing on defence-in-depth, their interaction is expected to explain cybersecurity risk management outcomes beyond their separate effects. Organisational security policy moderates these relationships, so that technically identical controls may produce different outcomes under different policy regimes. Risk management theory shapes the computational layer, in which the threat-level effectiveness measure, the firewall configuration score, and the NAC strength score combine into a composite institutional index. Design science frames the outer layer: the framework is benchmarked against its base models and validated through expert, Delphi, and scenario evidence before it is refined.

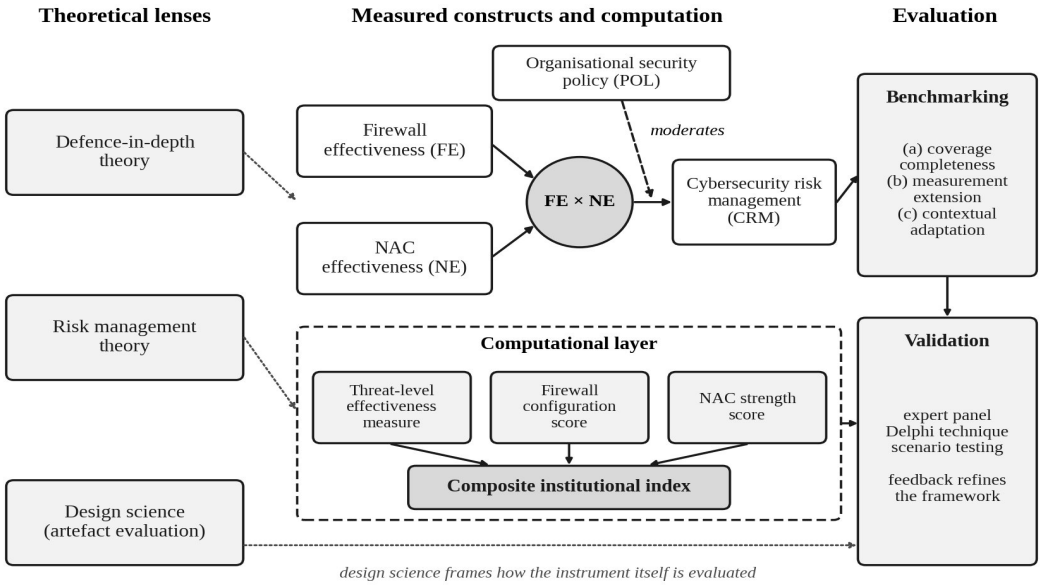


Figure 1: Conceptual Framework of the Study
Note: *FE = firewall effectiveness; NE = network access control effectiveness; POL = organisational security policy; CRM = cybersecurity risk management.*
Source: *Author-generated (2026).*

METHODOLOGY

This section explains the research design and its rationale, the sampling procedure, and how the framework will be implemented, benchmarked, validated, and analysed.

Research Design and Rationale for Mixed Methods

The study adopts a pragmatic research philosophy, which selects methods according to what each research question requires rather than by allegiance to a single paradigm (Morgan, 2007). Within that stance, it follows a design science approach for building and evaluating the framework (Hevner et al., 2004; Peffers et al., 2007) and a convergent mixed-methods design for gathering evidence (Creswell & Plano Clark, 2018).

Quantitative and qualitative methods are combined for three reasons. First, the study asks two different kinds of question. Measuring how well firewall and NAC controls perform, and how strongly they relate to risk management outcomes, requires standardised data that can be analysed statistically. Explaining why implementation gaps persist requires the judgement of the people who operate those controls, which interviews capture far better than questionnaires. Second, the strands check each other. Questionnaire responses are compared with documents and interview accounts, so that inflated self-reports can be detected rather than averaged away. Third, validating the framework depends on expert judgement, which is qualitative in nature and poorly served by survey data alone.

The two strands run in parallel and are integrated at interpretation. This convergent structure was chosen over a sequential one for a practical reason: access to banks cannot be assumed twice, so collecting both strands within the same access window is the safer arrangement.

Study Setting, Population, and Sampling

The study is set in Tier-One Nigerian commercial banks. These institutions run the most complex networks in the sector, carry the largest share of electronic transaction value, and employ dedicated security staff who can answer technical questions. The population comprises personnel in six categories whose duties involve network security controls: network administrators, cybersecurity officers, information technology (IT) auditors, information security managers, risk managers, and information and communication technology (ICT) managers. Other staff are excluded because they cannot answer questions on rule management or posture assessment.

Purposive sampling is applied to both strands, so that every respondent holds the technical knowledge the instrument presupposes. A comparable study of Nigerian deposit money banks used purposively selected senior staff for the same reason (Akintoye et al., 2022). The weakness of purposive sampling is that it does not support statistical generalisation. Three measures limit selection bias: respondents are drawn from every participating bank, all six job categories are represented, and the selection criteria are stated explicitly.

Banks do not publish the size of their security functions. The accessible population (N) will therefore be confirmed by formal request to each participating bank. The questionnaire sample (n) will then be computed with the Yamane (1967) formula, $n = N / (1 + Ne^2)$, at a 5% margin of error ($e = .05$). The result will be cross-checked against the tables of Krejcie and Morgan (1970), and questionnaires will be distributed approximately 20% above the computed minimum to absorb non-response. For the qualitative strand, 12 to 15 interviews will be conducted with senior information security and ICT managers. This range is consistent with

evidence that thematic saturation typically occurs within the first 12 interviews (Guest et al., 2006). A separate panel of 8 to 12 experts will validate the framework (Section 3.5.2). The questionnaire will be piloted with approximately 20 respondents drawn from the target population but excluded from the main sample.

Framework Implementation

Figure 2 summarises how the framework is implemented and validated. Implementation has four elements: the inputs, an integration and analysis layer, an evidence reconciliation step, and iterative development.

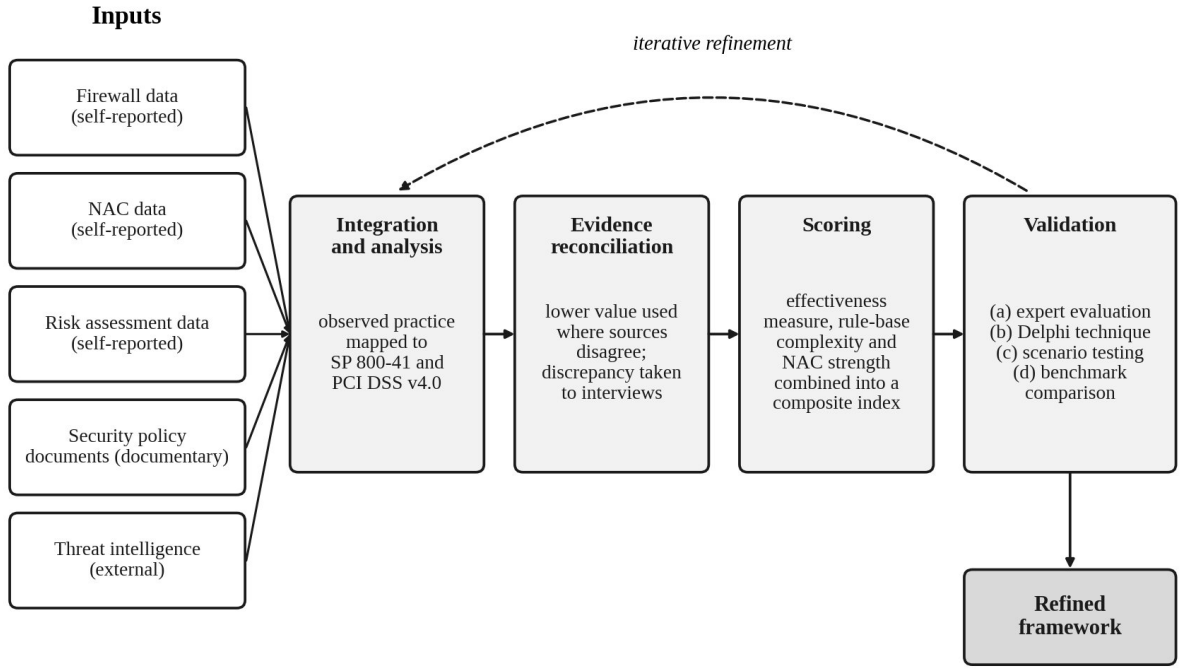


Figure 2: Implementation and Validation Workflow of the Composite Framework

Note: *SP 800-41 = NIST Special Publication 800-41 Revision 1; PCI DSS = Payment Card Industry Data Security Standard; NAC = network access control.*
Source: *Author-generated (2026).*

Inputs

Five categories of input feed the framework: (a) firewall data, (b) NAC data, (c) risk assessment data, (d) security policy documentation, and (e) threat intelligence. Firewall data comprise configuration practices, rule management arrangements, and the rule-base measures from which anomaly density, over-permissiveness, and rule utilisation are computed. NAC data comprise authentication mechanisms, posture assessment arrangements, and compliance enforcement practice. Risk assessment data describe the threat and vulnerability profile of the participating institution, while security policy documentation sets out the governance rules within which the framework must operate. Threat intelligence covers the attack techniques currently directed at financial institutions. Without this last input, a risk assessment is built against yesterday's threats, and the effectiveness scores inherit that staleness.

The inputs are deliberately heterogeneous. Three are self-reported, one is documentary, and one is external. That mixture is what makes the reconciliation step described in Section 3.3.3 both possible and necessary.

The Integration and Analysis Layer

Inputs pass through an integration layer structured by the base models. Observed practice is mapped onto the requirements of SP 800-41 and PCI DSS v4.0, which makes the distance between current practice and standard expectation explicit and measurable rather than impressionistic. The framework is then constructed to close that distance.

Four specifications follow from this. The framework states how firewall and NAC components should be configured to operate as a coordinated pair rather than as independent controls. It states how identity and posture information should inform traffic filtering decisions. It states how segmentation should be enforced across both layers. And it states how the mathematical components are computed and interpreted.

Evidence Reconciliation

Where questionnaire responses and documentary evidence disagree, the computation takes the lower value, flags the discrepancy, and carries it into the interview phase for explanation. This is a deliberate asymmetry. Self-reported security maturity tends to exceed observed maturity, because self-report measures are vulnerable to social desirability and related response biases (Podsakoff et al., 2003). An instrument that averages the two would conceal exactly the finding it should surface.

The reconciliation step is the point at which the framework departs most clearly from survey-based assessment. A purely self-reported instrument cannot detect the gap between what an institution believes its controls do and what its configuration shows them doing.

Iterative Development

Development proceeds iteratively rather than linearly. Feedback from the validation described in Section 3.5 is used to refine the framework, and the refined version is reassessed before the final form is presented. A framework that is validated once, at the end, has been tested rather than developed; the distinction matters for what the validation evidence can support.

Benchmarking Procedure

Benchmarking proceeds across five criteria: (a) coverage, (b) measurability, (c) integration, (d) contextual fit, and (e) validation evidence. Coverage refers to the control dimensions each base model specifies, and measurability to whether a criterion can be scored rather than merely attested. Integration asks whether firewall and NAC are evaluated as a coordinated pair. Contextual fit concerns the Nigerian regulatory environment, and validation evidence is the evidence that supports use. Three tests operationalise these criteria, and each has a stated pass condition (Table 3).

The first test, coverage completeness, requires every control criterion specified in SP 800-41 and in PCI DSS Requirements 1 and 7 to be represented within the framework. Any omission is treated as a defect requiring correction, not as a scope decision.

The second test, measurement extension, identifies the criteria for which the framework provides a quantified score where the base models provide only a binary compliance judgement. This is the principal claimed contribution and the place where the framework must demonstrably add something.

The third test, contextual adaptation, requires every criterion modified for the Nigerian setting to be documented together with the justification for the modification. Undocumented divergence from a base model is a failure of this test, not an adaptation.

The framework will be considered an improvement on its base models only where it demonstrates full coverage, adds measurement capability, and remains applicable within the local regulatory environment. Failing any one of the three is a negative result and will be reported as such.

Table 3: Benchmarking Tests and Pass Conditions

Test	Question addressed	Pass condition	Consequence of failure
Coverage completeness	Is every criterion in SP 800-41 and PCI DSS Requirements 1 and 7 represented?	Every criterion is mapped to at least one framework indicator	The omission is treated as a defect and corrected
Measurement extension	Where does the framework score what the base models only attest?	Criteria judged only as pass or fail by the base models receive a quantified score	No demonstrated contribution; reported as a negative result
Contextual adaptation	Is every modification for the Nigerian setting documented?	Each modified criterion is documented with its justification	Undocumented divergence counts as failure

Note: *SP 800-41 = NIST Special Publication 800-41 Revision 1; PCI DSS = Payment Card Industry Data Security Standard.*
Source: *Author-generated (2026), with base-model criteria from Scarfone and Hoffman (2009) and PCI SSC (2022).*

Validation Design

Strategy

Validation establishes whether the framework performs the function for which it was constructed. A multi-method strategy is applied, following guidance that evaluation in design research should be planned deliberately with respect to why, when, how, and what is evaluated (Venable et al., 2016).

No single method is sufficient here. Expert judgement establishes whether the instrument is usable and complete but cannot establish whether it detects what it should. Scenario testing establishes detection but not practicality. Benchmarking establishes relative contribution but says nothing about either. The four methods described below, summarised in Table 4, are chosen to cover one another's blind spots.

Table 4: Validation Methods and What Each Establishes

Method	What it establishes	What it cannot establish	Evidence produced
Expert evaluation (8 to 12 panellists)	Usability, completeness, and practicality	Whether the framework detects real weaknesses	Ratings against five criteria; content validity ratio for each item
Delphi technique	Convergence on contested judgements, chiefly the weights	Whether the agreed position is correct	Second-round ratings; documented disagreement
Scenario-based validation	Detection of the weakness each documented attack exploits	Practicality in routine use	Detection outcome for each scenario
Benchmark comparison	Contribution relative to the base models	Usability or detection	Results of the three tests in Table 3

Source: *Author-generated (2026), summarising the validation design in Sections 3.5.1 to 3.5.5, informed by Venable et al. (2016).*

Expert Evaluation

Expert evaluation is the primary method. A panel of 8 to 12 practitioners and academics holding demonstrable experience in network security, banking technology, or information security governance assesses the framework against defined criteria using a structured instrument. Content validity of that instrument follows the quantitative approach of Lawshe (1975), with a content validity ratio computed for each item.

Five criteria govern the evaluation: (a) usability, (b) effectiveness, (c) reliability, (d) scalability, and (e) practicality. Usability asks whether practitioners can understand and apply the framework. Effectiveness asks whether it identifies the control weaknesses it was designed to detect. Reliability asks whether it produces consistent scores across assessors and conditions, and scalability asks whether it remains workable as network size and transaction volume increase. Practicality asks whether it can be applied within the budgetary, staffing, and regulatory constraints that a Nigerian commercial bank actually faces. The last criterion is the one most often omitted from framework validation and the one most likely to decide whether an instrument is used.

The Delphi Technique

Where the first round of expert assessment produces materially divergent judgements, the Delphi technique is applied. Anonymised summaries of panel responses are returned to participants for a second round, permitting convergence towards a considered group position without the dominance effects that face-to-face deliberation introduces (Dalkey & Helmer, 1963; Skulmoski et al., 2007). Rounds continue until convergence or until it becomes clear that a genuine disagreement exists, which is itself a reportable finding rather than a problem to be smoothed away.

The weights are the element most exposed to expert challenge. Those proposed for the rule-base complexity index and for the composite index are reasoned rather than derived, and Delphi elicitation is intended to refine them. Sensitivity analysis across plausible weight ranges will establish whether the composite ordering of institutions is robust to weight choice, since that is the property that matters for comparative use.

Scenario-Based Validation

Scenario-based validation tests the framework against documented attack patterns, including unauthorised device connection, insider misuse of legitimate credentials, lateral movement following endpoint compromise, and unauthorised external access. For each scenario the framework's computed output is examined to determine whether it would have identified the control weakness the scenario exploits.

This is the sharpest test in the design, because it can fail cleanly. A framework that returns a strong score for an institution whose controls would not have stopped the scenario has been shown to measure the wrong thing.

Benchmark Comparison

The benchmark comparison applies the three tests specified in Section 3.4: coverage completeness, measurement extension, and contextual adaptation. Results are reported test by test rather than as an overall verdict, so that partial success is visible as partial.

Analytical Methods and Software

Quantitative Analysis

Questionnaire data are prepared through screening for incomplete responses, coding, and testing of distributional assumptions. Descriptive statistics summarise respondent characteristics and observed levels of firewall and NAC effectiveness. Internal consistency is assessed through Cronbach's alpha, computed separately for each construct scale rather than for the instrument as a whole, since a single overall coefficient can conceal a weak subscale (Cronbach, 1951).

Inferential analysis proceeds in three stages. Correlation analysis establishes the direction and strength of association between firewall effectiveness, NAC effectiveness, and cybersecurity risk management. Multiple regression then estimates the extent to which the two independent variables jointly predict risk management outcomes, taking the form shown in Equation 1:

$$\mathrm{CRM} = b_0 + b_1(\mathrm{FE}) + b_2(\mathrm{NE}) + b_3(\mathrm{FE} \times \mathrm{NE}) + b_4(\mathrm{POL}) + e \quad (1)$$

where CRM denotes cybersecurity risk management, FE denotes firewall effectiveness, NE denotes NAC effectiveness, FE × NE denotes their interaction, POL denotes organisational security policy, and e denotes the error term.

The interaction term carries the weight of the argument. A significant coefficient on b_3 would constitute statistical evidence that firewall and NAC effectiveness combine rather than simply add, which is the integration claim the framework is built on. A non-significant coefficient would be a negative result for that claim and would need to be reported as one.

Structural equation modelling then tests the full measurement and structural model, including the moderating effect of organisational security policy. Partial least squares (PLS) estimation

is preferred to covariance-based alternatives on three grounds: it accommodates smaller samples, imposes no requirement of multivariate normality, and suits research that is predictive and exploratory rather than confirmatory of settled theory (Hair et al., 2019; Shiau et al., 2019). Measurement model quality is assessed through composite reliability, average variance extracted (Fornell & Larcker, 1981), and the heterotrait-monotrait ratio of correlations, the last preferred because it detects a lack of discriminant validity in variance-based models that the older criteria miss (Henseler et al., 2015). Where the measurement model proves consistent and reflective throughout, the consistent PLS estimator is reported alongside the conventional one, so that any divergence between the two is visible rather than concealed by the choice of algorithm (Dijkstra & Henseler, 2015).

Common method bias is a live threat, since independent and dependent constructs are measured from the same respondents at one point in time. Procedural remedies are applied, including separation of construct blocks, assurance of anonymity, and avoidance of leading item wording (Podsakoff et al., 2003).

Descriptive statistics, correlation analysis, and the multiple regression in Equation 1 will be conducted in IBM SPSS Statistics (Version 29). The PLS structural equation model will be estimated in SmartPLS (Version 4).

Qualitative Analysis

Interview data are analysed thematically through familiarisation, initial coding, searching for themes, reviewing themes, defining and naming themes, and reporting (Braun & Clarke, 2006, 2022). Coding decisions are documented in a codebook maintained throughout, following display and verification procedures that keep the route from transcript to reported theme auditable (Miles et al., 2014). Where interview material concerns matters of fact about configuration rather than interpretation, directed content analysis is applied instead, using categories derived from the base models (Hsieh & Shannon, 2005). Coding will be managed in NVivo (Version 14).

Integration

Integration occurs at interpretation, where statistical findings and interview themes are set against one another. Convergence strengthens confidence in a finding. Divergence is treated as informative and reported rather than reconciled, since it frequently marks the distance between formal practice and operational reality, which is precisely the gap the framework exists to measure.

FINDINGS: DESIGN OUTCOMES

Because the design has been specified but not yet executed, the findings reported here are design outcomes rather than empirical results. In design science terms, they constitute an ex ante evaluation of the framework, established before field data are collected (Venable et al., 2016). Five outcomes follow.

The Framework Traces to Documented Requirements

Table 5 shows that every firewall and NAC dimension in the framework maps to a documented requirement in a base model or a supporting standard. Firewall dimensions trace to SP 800-41 and to PCI DSS Requirement 1. NAC dimensions trace to PCI DSS Requirement 7 and to NIST

SP 800-207. This traceability, built on the base models compared in Table 2, is the precondition for the coverage completeness test. It means that any gap the test later reveals will appear as an unmapped criterion rather than being hidden inside a composite score.

Table 5: Traceability of Framework Dimensions to Base-Model Requirements

Construct	Dimension	Indicator	Source requirement
FE	Traffic filtering	Threat detection rate; malicious traffic blocked	SP 800-41, Section 4
FE	Policy management	Rule accuracy; review frequency; default-deny posture	SP 800-41, Section 5; PCI DSS Requirement 1
FE	Rule-base integrity	Shadowed, redundant, and correlated rules	SP 800-41, Section 5
FE	Inspection capability	Deep packet inspection coverage; encrypted traffic visibility	SP 800-41, Section 2
NE	Authentication	User and device verification; 802.1X coverage	SP 800-207
NE	Authorisation	Least-privilege assignment; role granularity	PCI DSS Requirement 7
NE	Posture assessment	Patch, endpoint protection, and encryption checks at admission	SP 800-207
NE	Compliance enforcement	Device compliance rate; quarantine of non-compliant devices	SP 800-207

Note: *FE = firewall effectiveness; NE = network access control effectiveness; SP = NIST Special Publication; PCI DSS = Payment Card Industry Data Security Standard.* **Source:** *Author-generated (2026), mapping indicators to Scarfone and Hoffman (2009), PCI SSC (2022), and Rose et al. (2020).*

Benchmarking is Reduced to Falsifiable Tests

Table 3 sets out the second outcome. Benchmarking is reduced to three tests, each with a pass condition and a stated consequence of failure. The framework can therefore fail benchmarking in a specific, reportable way, rather than passing by general resemblance to its base models.

The Validation Methods are Complementary

Table 4 and Figure 2 show the third outcome. The four validation methods are complementary by construction, because each establishes something the others cannot. Expert evaluation tests usability and completeness, scenario testing tests detection, and benchmarking tests relative contribution. The Delphi technique, in turn, resolves divergent judgements, chiefly about the weights. The iterative loop in Figure 2 returns validation feedback to the framework before its final form is fixed.

Evidence Reconciliation Is the Principal Methodological Departure

The fourth outcome is the reconciliation rule described in Section 3.3.3. Where self-report and documentary evidence disagree, the lower value is used and the discrepancy is carried into the interviews. This rule distinguishes the framework from survey-based assessment, which cannot detect the gap between what an institution believes its controls do and what its configuration shows.

The Integration Claim Becomes Testable

The fifth outcome links the theoretical and computational layers of Figure 1. The defence-in-depth proposition that firewall and NAC are complementary is expressed as the interaction term b_3 in Equation 1. A significant coefficient would support the integration claim. A non-significant one would count against it, and it would be reported as such.

DISCUSSION

The design outcomes can now be set against the literature reviewed in Section 2.2. The comparison shows where the framework supports existing work, where it challenges it, and where it extends it.

Where the Design Supports Existing Work

The decision to pair the firewall with NAC is consistent with firewall research. Heino et al. (2022) conclude that no single endpoint-identification method is sufficient, which implies that traffic inspection needs an admission control alongside it. The emphasis on dynamic verification of identity and device posture also aligns with Zero Trust research (Gambo & Almulhem, 2026; Rose et al., 2020). Likewise, scoring configuration and access policy follows established precedent (Belhaouane et al., 2015; Bradatsch et al., 2023; Scarfone & Mell, 2010).

Where the Design Challenges Existing Work

The design departs from the prevailing standard of evidence. Recent Zero Trust frameworks for financial institutions have been evaluated through demonstration applications, prototypes, or algorithm designs (Chaudhry & Hydros, 2023; Daah et al., 2023, 2024). By contrast, the present design requires validation in operating banks and benchmarking against recognised standards before any claim of improvement is made. It also addresses a limitation of survey-based studies of bank cybersecurity, which rely on self-report (Akintoye et al., 2022). The reconciliation rule treats inflated self-report as a finding to be surfaced, consistent with long-standing warnings about response bias in self-report data (Podsakoff et al., 2003).

Where the Design Extends Existing Work

Three extensions stand out. First, the framework narrows the institution-wide approach of Azura et al. (2025) to the firewall and NAC pair, and it recalibrates that pair to Nigerian regulation. Second, it complements financial quantification models such as that of Pollmeier et al. (2023) by tying scores to configuration evidence rather than to loss estimates alone. Third, it responds to the fragmentation that Sopko and Šafár (2026) identify, because it connects a technical measurement instrument to the governance and regulatory requirements of the CBN (2024) framework.

Contribution to Knowledge

The study's contribution is therefore threefold, and it answers the four gaps summarised in Table 1. Methodologically, it shows how a security evaluation framework can be benchmarked through falsifiable tests and validated through complementary methods specified in advance. Theoretically, it converts the defence-in-depth claim of complementarity into a testable hypothesis about interaction. Contextually, it calibrates a firewall and NAC instrument to the regulatory environment of Nigerian deposit money banks, a setting that the international literature rarely reaches.

Implications

Implications for Research

Three implications follow for research. First, evaluation frameworks in cybersecurity should state their validation criteria before execution. The benchmarking tests in Table 3 offer a template that other researchers can adapt when they validate their own instruments. Second, the complementarity of security controls should be tested rather than assumed. The interaction term in Equation 1 provides one way to do so, and it can be applied to other control pairs as Zero Trust architectures mature (Gambo & Almulhem, 2026). Third, researchers who rely on self-reported security data should treat the gap between self-report and documentary evidence as an object of study in its own right (Podsakoff et al., 2003).

Implications for Practice

For banks, the framework offers a way to evidence control effectiveness under the self-assessment that the CBN (2024) framework requires, using scores rather than attestation. For regulators, the traceability shown in Table 5 suggests how supervisory reviews could move from checking whether controls exist to measuring how well they perform. For security practitioners, the reconciliation rule and the rule-base measures point to specific, improvable weaknesses, such as shadowed rules and incomplete posture checks, rather than to general maturity levels. Finally, because the framework expresses effectiveness as risk reduction, it gives boards a defensible basis for prioritising security investment (ISO, 2018).

Limitations

Five limitations are acknowledged, and they bound what the design can deliver.

Execution Status: The design is specified but not executed. Every claim in this paper concerns construction rather than demonstrated performance, and the paper should be read as a protocol.

Weight Derivation: The weights in the complexity index and the composite index are reasoned rather than empirically derived. The sensitivity analysis described in Section 3.5.3 is intended to establish how much this matters. Until it is run, however, the composite score carries an unquantified dependence on choices made by the designer.

Access to Rule-Base Exports: The firewall component depends on access to rule-base exports, which institutional security policy may preclude. Where that access is refused, component measures are obtained through respondent estimation and reported as estimated rather than measured. Values obtained that way are not comparable with measured values, and mixing them in a single analysis would be a mistake.

Panel Composition: Expert panel validation is vulnerable to the composition of the panel. A panel of 8 to 12 practitioners drawn from one national banking sector may share assumptions that a wider panel would challenge.

Context and Scale: Finally, the framework is calibrated to the Nigerian regulatory environment and to institutions of a particular scale. Application elsewhere would require recalibration of the threat criticality weights at minimum, and possibly of the dimension weights.

Future Research Directions

Several directions follow from this protocol. The most immediate is execution: applying the framework in selected Tier-One Nigerian banks and reporting the results against the criteria set out here, including any negative results.

A second direction concerns the weights. Future work can combine Delphi elicitation with sensitivity analysis to derive the weights empirically and to test whether institutional rankings remain stable under alternative weightings.

Third, repeated application over time would show whether the framework detects improvement or decline, which is the property that matters most for supervisory use. Longitudinal designs would also reveal whether the gap between self-reported and documented maturity narrows once it is measured.

Fourth, the framework should be tested beyond its original setting. Tier-Two banks, microfinance institutions, payment service banks, and financial institutions in other African markets operate under different resource and regulatory constraints. Recalibrating the framework for them would test its generalisability.

Finally, future research could extend the framework to additional controls, such as intrusion detection and security information and event management, and examine how its measures change as institutions move towards Zero Trust architectures (Rose et al., 2020).

CONCLUSION

This paper has specified how a composite framework for evaluating firewall and NAC effectiveness will be implemented, validated, and benchmarked. Implementation is a five-input pipeline passing through an integration layer structured by two technical base models, with an evidence reconciliation step that resolves self-report against documentation in favour of the lower value. Validation combines expert panel assessment, the Delphi technique, scenario-based testing, and benchmark comparison, each chosen to cover the blind spots of the others. Benchmarking is reduced to three tests with stated pass conditions, so that the framework can fail them.

The contribution of the design lies in its falsifiability. A framework benchmarked on general similarity to a standard cannot be shown to be wrong; one required to demonstrate coverage completeness, measurement extension, and documented contextual adaptation can be. Specifying the criteria before execution is what makes the eventual result meaningful either way. Execution follows: panel recruitment, instrument piloting, and application in selected Nigerian Tier-One commercial banks. The results, including any negative ones, will be reported against the criteria set out here.

The larger lesson reaches beyond firewalls and NAC. Security evaluation earns trust when it can be proved wrong. By fixing its criteria before execution, tying every score to documented requirements, and treating inflated self-report as evidence rather than noise, the framework offers banks and regulators a practical route from attestation to measurement. It also gives researchers a protocol against which future evaluation frameworks can be built and judged.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

Declaration of Generative AI Use

During the preparation of this manuscript, the authors used Claude (Anthropic) to assist with drafting, language editing, formatting, and reference checking. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

REFERENCES

Akintoye, R., Ogunode, O., Ajayi, M., & Joshua, A. A. (2022). Cyber security and financial innovation of selected deposit money banks in Nigeria. *Universal Journal of Accounting and Finance, 10*(3), 643–652. https://doi.org/10.13189/ujaf.2022.100302

Al-Shaer, E. S., & Hamed, H. H. (2004). Discovery of policy anomalies in distributed firewalls. In *IEEE INFOCOM 2004* (Vol. 4, pp. 2605–2616). IEEE. https://doi.org/10.1109/INFCOM.2004.1354680

Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research, 253*(1), 1–13. https://doi.org/10.1016/j.ejor.2015.12.023

Azura, Y. T. Y., Azad, M. A., & Ahmed, Y. (2025). An integrated cyber security risk management framework for online banking systems. *Journal of Banking and Financial Technology, 9*(1), 85–104. https://doi.org/10.1007/s42786-025-00056-3

Belhaouane, M., Garcia-Alfaro, J., & Debar, H. (2015). Evaluating the comprehensive complexity of authorization-based access control policies using quantitative metrics. In *Proceedings of the 12th International Conference on Security and Cryptography* (pp. 53–64). SCITEPRESS. https://doi.org/10.5220/0005544100530064

Boungou, W. (2023). Cyber-attacks and banking intermediation. *Economics Letters, 233*, Article 111354. https://doi.org/10.1016/j.econlet.2023.111354

Bradatsch, L., Miroshkin, O., Trkulja, N., & Kargl, F. (2023). Zero trust score-based network-level access control in enterprise networks. In *2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 1422–1429). IEEE. https://doi.org/10.1109/TrustCom60117.2023.00194

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology, 9*(1), 3–26. https://doi.org/10.1037/qup0000196

Central Bank of Nigeria. (2024). *Risk-based cybersecurity framework and guidelines for deposit money banks and payment service banks*. https://www.cbn.gov.ng/Out/2024/BSD/CBN%20Risk-Based%20Cybersecurity%20Framework%20for%20DMBs%20and%20PSBs_2024.pdf

Chaudhry, U. B., & Hydros, A. K. M. (2023). Zero-trust-based security model against data breaches in the banking sector: A blockchain consensus algorithm. *IET Blockchain, 3*(2), 98–115. https://doi.org/10.1049/blc2.12028

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Daah, C., Qureshi, A., & Awan, I. (2023). Zero trust model implementation considerations in financial institutions: A proposed framework. In *2023 10th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 71–77). IEEE. https://doi.org/10.1109/FiCloud58648.2023.00019

Daah, C., Qureshi, A., Awan, I., & Konur, S. (2024). Enhancing zero trust models in the financial industry through blockchain integration: A proposed framework. *Electronics, 13*(5), Article 865. https://doi.org/10.3390/electronics13050865

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9*(3), 458–467. https://doi.org/10.1287/mnsc.9.3.458

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fuszder, M. H. R., Abdullah, M., Sulong, Z., & Abakah, E. J. A. (2026). Cybersecurity risk and bank competition. *Journal of International Financial Markets, Institutions and Money, 109*, Article 102313. https://doi.org/10.1016/j.intfin.2026.102313

Gambo, M. L., & Almulhem, A. (2026). Zero trust architecture: A systematic literature review. *Journal of Network and Systems Management, 34*(1), Article 25. https://doi.org/10.1007/s10922-025-09998-x

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Gyamfi, B. A., Arhinful, R., & Mensah, L. (2026). Financial technology adoption and its influence on the cybersecurity performance of financial institutions listed on the Tokyo Stock Exchange. *Computers & Security, 171*, Article 105127. https://doi.org/10.1016/j.cose.2026.105127

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hassan, A. O., Ewuga, S. K., Abdul, A. A., Abrahams, T. O., Oladeinde, M., & Dawodu, S. O. (2024). Cybersecurity in banking: A global perspective with a focus on Nigerian practices. *Computer Science & IT Research Journal, 5*(1), 41–59. https://doi.org/10.51594/csitrj.v5i1.701

Heino, J., Hakkala, A., & Virtanen, S. (2022). Study of methods for endpoint aware inspection in a next generation firewall. *Cybersecurity, 5*(1), Article 25. https://doi.org/10.1186/s42400-022-00127-8

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

IBM. (2024). *Cost of a data breach report 2024*. https://www.ibm.com/reports/data-breach

International Organization for Standardization. (2018). *Risk management: Guidelines* (ISO Standard No. 31000:2018). https://www.iso.org/standard/65694.html

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607–610. https://doi.org/10.1177/001316447003000308

Kuipers, D., & Fabro, M. (2006). *Control systems cyber security: Defense in depth strategies*. Idaho National Laboratory. https://doi.org/10.2172/911553

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Leszczyna, R. (2021). Review of cybersecurity assessment methods: Applicability perspective. *Computers & Security, 108*, Article 102376. https://doi.org/10.1016/j.cose.2021.102376

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Onyeacholem, A. O., & Sanusi, F. A. (2026). *A composite index for evaluating firewall and network access control effectiveness in financial institutions* [Manuscript submitted for publication]. Department of Computer Science, Babcock University.

Osman, M. (2026). Governing the human-machine identity continuum: An integrated framework for cybersecurity risk management in financial services. *Computers & Security, 170*, Article 105038. https://doi.org/10.1016/j.cose.2026.105038

PCI Security Standards Council. (2022). *Payment Card Industry Data Security Standard: Requirements and testing procedures* (Version 4.0). https://www.pcisecuritystandards.org/document_library/

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

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), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Pollmeier, S., Bongiovanni, I., & Slapničar, S. (2023). Designing a financial quantification model for cyber risk: A case study in a bank. *Safety Science, 159*, Article 106022. https://doi.org/10.1016/j.ssci.2022.106022

Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

Scarfone, K., & Hoffman, P. (2009). *Guidelines on firewalls and firewall policy* (NIST Special Publication 800-41, Rev. 1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-41r1

Scarfone, K., & Mell, P. (2010). *The common configuration scoring system (CCSS): Metrics for software security configuration vulnerabilities* (NIST Interagency Report 7502). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.7502

Shiau, W.-L., Sarstedt, M., & Hair, J. F. (2019). Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Research, 29*(3), 398–406. https://doi.org/10.1108/IntR-10-2018-0447

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research, 6*, 1–21. https://doi.org/10.28945/199

Sopko, J., & Šafár, L. (2026). Cybersecurity and financial systems: A global perspective on research fragmentation and innovation gaps. *SN Business & Economics, 6*(4), Article 90. https://doi.org/10.1007/s43546-026-01105-9

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A framework for evaluation in design science research. *European Journal of Information Systems, 25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

Wang, S., Asif, M., Shahzad, M. F., & Ashfaq, M. (2024). Data privacy and cybersecurity challenges in the digital transformation of the banking sector. *Computers & Security, 147*, Article 104051. https://doi.org/10.1016/j.cose.2024.104051

Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper & Row.

APPENDIX

List of Acronyms

802.1X	Institute of Electrical and Electronics Engineers standard for port-based network access control
AI	Artificial intelligence
CBN	Central Bank of Nigeria
CRM	Cybersecurity risk management
FE	Firewall effectiveness
FEDS	Framework for Evaluation in Design Science
IBM	International Business Machines Corporation
ICT	Information and communication technology
ISO	International Organization for Standardization
IT	Information technology
NAC	Network access control

NDPC	Nigeria Data Protection Commission
NE	Network access control (NAC) effectiveness
NIST	National Institute of Standards and Technology
ORCID	Open Researcher and Contributor ID
PCI DSS	Payment Card Industry Data Security Standard
PCI SSC	Payment Card Industry Security Standards Council
PLS	Partial least squares
POL	Organisational security policy
SP	Special Publication (NIST)
SPSS	Statistical Package for the Social Sciences (IBM SPSS Statistics)
USD	United States dollar