

Business Intelligence Dashboards for Household Financial Planning Tax Management and Decision Making

Valentina Ochuko Obukadata^{1*}, Ifeoma E. Okoli², Oluwatobiloba Mide-Ibrahim³,
Leonard U. Agwuna⁴, Rosita Ebere Daraojimba⁵

¹ University of New Haven, West Haven, Connecticut, USA

² Salesforce Development Company, Canada

³ FinServe Pro, USA

⁴ School of Commerce, University of KwaZulu-Natal, Durban, South Africa

⁵ First Bank of Nigeria Plc, Lagos, Nigeria

* Corresponding author: Valentina Ochuko Obukadata, obukadataval@gmail.com

Abstract

Purpose: Personal finance applications now aggregate bank, card, loan and tax data into dashboards that promise better household decisions, yet their evaluation rarely goes beyond engagement and satisfaction. This review asks which dashboard design properties improve the quality of household cash-flow, debt, saving and tax decisions, and under which conditions they produce over-reliance or privacy harm. **Method:** An integrative review of household finance, behavioral economics, visual analytics, fiscal psychology and data governance research published to 2025 was conducted under a documented search, screening and appraisal protocol, with contradictory findings adjudicated by design strength and outcome proximity. **Findings:** Evidence converges on four conditions: reconciled data, perceptually efficient encodings matched to the user's visualization literacy, explicit uncertainty on forecasts and tax estimates, and prompts placed at the moment of choice. Evidence diverges on whether richer displays or simple rules of thumb serve low-literacy households better. **Contribution:** The article proposes the LEDGER framework (linked data, encoding fit, decision anchoring, guided action, exposed uncertainty, rights over data), defines decision-quality metrics including dominance errors and Brier-scored shortfall judgments, states six falsifiable propositions, and specifies a two-stage validation design. **Implications:** Providers and regulators should judge household dashboards by decision quality and harm incidence rather than logins or feature counts.

Keywords: business intelligence dashboards, household finance, open banking, visualization literacy, present bias, mental accounting, tax readiness, decision support

DOI: 10.18091/ijsts.v9i3.112

1 Introduction

Households make a small number of financial decisions that matter a great deal and a large number that matter a little, and they make most of both with incomplete information about their own position. A family deciding whether to accelerate a credit card repayment, raise a paycheck withholding election, or move money into an emergency fund must first know what it owns, what it owes, what is coming due, and what the tax authority will eventually expect. Business intelligence techniques that were developed for firms, namely integration of transactional sources, dimensional aggregation, key performance indicators and interactive visualization, have migrated into consumer products that promise exactly this overview (Mbonu et al., 2021a; Sanni et al., 2021b). The migration has been accelerated by open banking regimes that let a consumer authorize a third party to read account data, and by tax preparation software that now offers year-round projections rather than a single filing-season interaction (Bola-Sadipe et al., 2022; Oluwo et al., 2025).

The premise of these products is that visibility improves choice. That premise is plausible but not self-evident. Household finance research documents persistent gaps in basic numeracy, debt literacy and understanding of interest compounding, and those gaps are associated with costly borrowing and weak saving (Lusardi et al., 2015; Atkinson et al., 2012). A dashboard that presents an annual percentage rate, a net-worth sparkline and a tax liability estimate to a user who cannot interpret them may add confidence without adding understanding. Behavioral research adds a second complication: people partition money into mental accounts, weight the present disproportionately, and respond to salience and defaults, so the same information laid out differently can produce different decisions (Fobellah, 2025a; Xiao et al., 2016). A third complication is that aggregation requires the household to share credentials or grant access tokens covering its most sensitive records, which introduces privacy and security risks that the decision benefit must outweigh (Kumuyi et al., 2024b; Mbonu et al., 2022a).

1.1 Research Gap and Questions

Three strands of work bear on household dashboards but rarely meet. Visual analytics and business intelligence research studies how displays support monitoring and performance management, mostly in organizational settings with trained analysts and defined accountabilities (Ogbole et al., 2021; Tonoyan et al., 2024b). Household finance and financial education research measures literacy, capability and well-being, but treats information delivery as a course or a disclosure rather than as an interactive display (Miller et al., 2015; Netemeyer et al., 2018). Tax compliance research explains why people pay, underpay or overpay, with attention to trust, enforcement and complexity, but seldom examines the interface through which a taxpayer forms an expectation of liability (Alm, 2019; Kirchler, 2007). The gap is therefore not a shortage of studies on any one topic; it is the absence of a framework that connects dashboard design properties to the quality of specific household decisions, identifies when those properties backfire, and states what evidence would show that a dashboard helped. Accordingly, the review addresses three research questions. RQ1: Which design properties of household financial dashboards are associated with more accurate comprehension and better decisions about cash flow, debt, saving and tax? RQ2: Under what household, data and design conditions do those properties fail or produce harm, including over-reliance on estimates and privacy loss? RQ3: How should decision quality be measured so that competing dashboard designs can be compared and falsified (Hevner et al., 2004; Wieringa, 2014)?

1.2 Contributions and Structure

The article makes four contributions. First, it synthesizes evidence across household finance, behavioral economics, graphical perception and fiscal psychology into a single account of how a dashboard can change a household decision, and it adjudicates rather than lists the places where that evidence conflicts. Second, it proposes the LEDGER framework, whose six constructs (linked and reconciled data, encoding fit, decision anchoring, guided action, exposed uncertainty and rights over data) name design properties that can be manipulated experimentally. Third, it defines decision-quality metrics that go beyond satisfaction and usage: dominance error rates, the expected cost of decision error, Brier scores for probabilistic shortfall judgments, and tax estimate error relative to safe-harbor thresholds. Fourth, it states six propositions with explicit tests and outlines a two-stage validation design, pairing a randomized decision experiment with a stepped-wedge field evaluation (Easterbrook et al., 2008; Runeson et al., 2009). Section 2 sets out the theoretical foundations. Section 3 documents the review protocol. Section 4 presents the synthesis. Section 5 develops the framework, metrics, a worked scenario and the propositions. Section 6 describes the validation design and threats to validity. Section 7 discusses implications, counter-arguments, limitations and a research agenda, and Section 8 concludes.

2 Theoretical Foundations

Four bodies of theory explain why a household dashboard might help and why it might not. Bounded rationality and its behavioral extensions explain the systematic departures from optimal choice that a display could correct or amplify. Graphical perception and visualization literacy explain whether the display is decoded accurately. Fiscal psychology explains how taxpayers form and act on beliefs about liability. Theories of reliance and planned behavior explain when an accurate belief becomes an action. The sections below use each theory to derive a specific expectation about design, rather than invoking it as general background, because a framework that cannot be traced to mechanism cannot be falsified in a meaningful way (Ajzen, 1991; Fobellah, 2025a; Kirchler, 2007).

2.1 Bounded Rationality, Mental Accounting and Present Bias

Standard intertemporal models assume that a household discounts future consumption exponentially, so that preferences between two dated options do not reverse as the earlier option approaches. Observed behavior departs from this in a structured way. The quasi-hyperbolic formulation captures the departure: utility at time t is $U_t = u(c_t) + \beta \sum_{k \geq 1} \delta^k u(c_{t+k})$, where δ is the long-run discount factor and $\beta < 1$ introduces an extra penalty on every future period relative to the present. A present-biased household plans to save next month and then, when next month arrives, defers again. The implication for design is specific: displays that make a future cost vivid at the moment of a present choice, or that let the household commit in advance, should outperform displays that simply report balances (Fobellah, 2025a; Abah et al., 2025).

Mental accounting describes a second regularity: money is not treated as fungible. Households label funds as rent, holiday or tax money, track categories against budgets, and resist moving money between labels even when doing so would reduce total cost, as when a family holds low-yield savings while revolving a high-interest card balance. Dashboards interact with this tendency in two directions. Category budgets and savings goals can harness labeling to protect an emergency buffer or a tax reserve, which is useful when self-control is the binding constraint. The same labels can entrench non-fungible choices that a

consolidated view would expose. Which effect dominates depends on whether the display foregrounds the cross-account trade-off or hides it inside separate tiles (Drexler et al., 2014; Xiao et al., 2016). Present bias and mental accounting therefore generate opposite predictions for the same feature, which is why the framework treats decision anchoring as a construct to be measured rather than a feature to be added (Michael et al., 2021d).

2.2 Graphical Perception and Visualization Literacy

Research on graphical perception has established a broadly stable ranking of how accurately people decode quantitative encodings: position along a common scale is decoded most accurately, followed by length, then angle and slope, then area, with color saturation and hue least accurate for magnitude. Consumer finance applications frequently invert this ranking, favoring donut charts for spending shares, gauges for credit scores and colored tiles for budget status, because those encodings are compact and attractive on a phone screen. A bar chart of category spending against budget, sorted by overspend, is a position encoding; a donut chart of the same data asks the user to compare angles and arcs, and the error rate in magnitude judgments rises accordingly (Ogbole et al., 2021; Sanni et al., 2021b).

Visualization literacy, the ability to read, interpret and draw inferences from charts, varies widely and is only partly correlated with numeracy. Standardized assessment instruments ask respondents to retrieve values, find extrema, identify trends and compare groups across chart types, and they show that line charts and bar charts are read more reliably than stacked areas, bubble charts or dual-axis plots. For household dashboards the corollary is that encoding choice and user literacy interact: a fan chart that conveys forecast uncertainty to one user may be read by another as a range of equally likely outcomes, or ignored in favor of the central line. Measures of financial literacy such as those piloted across countries by the OECD capture knowledge of interest, inflation and risk diversification, but they do not capture chart comprehension, so the two must be measured separately (Atkinson et al., 2012; Tonoyan et al., 2024b).

2.3 Fiscal Psychology and Household Tax Behavior

Tax behavior is poorly explained by deterrence alone. Economic psychology describes compliance as the product of enforced compliance, driven by the perceived power of the authority, and voluntary cooperation, driven by trust; the slippery slope framework predicts that the two combine, and that interactions which build trust raise voluntary compliance (Kirchler, 2007). Reviews of compliance motives add social norms, tax morale, perceived fairness and, importantly for this review, complexity and uncertainty about what is owed (Alm, 2019). Many household tax errors are not evasion but misestimation: a worker with a second job whose withholding is calibrated to a single employer, a gig earner who does not set aside self-employment tax, or a family surprised by the loss of a credit when income crosses a phase-out threshold (Bola-Sadipe et al., 2022; Fobellah, 2025c).

For dashboard design, fiscal psychology implies that tax views serve two functions. The informational function reduces misestimation by projecting liability from year-to-date income, withholding and estimated payments. The relational function shapes how the household experiences the tax system: a projection that surfaces a looming balance due with a clear, feasible path to avoid a penalty frames the authority as predictable, whereas a projection that appears only after filing frames it as arbitrary. Loss framing also matters, since households react more strongly to a balance due than to an equivalent reduction in a refund, and many deliberately over-withhold to secure a refund that functions as a forced saving device. A dashboard that recommends precise withholding may therefore remove a commitment device the household valued, which is a design trade-off rather than an error (Kirchler, 2007; Anioke et al., 2023a).

2.4 Reliance, Trust Calibration and Planned Behavior

A dashboard changes a decision only if the user relies on it appropriately. Research on decision support in clinical and operational settings shows two symmetric failures: under-reliance, where accurate advice is ignored, and over-reliance, where inaccurate advice is followed because it comes from a system (Ilodigwe et al., 2025a; Adelanwa et al., 2024). Appropriate reliance requires that trust track the actual reliability of each output, which in turn requires that the display communicate reliability. A cash-flow forecast shown as a single line invites the user to treat it as a guarantee; the same forecast shown with an interval and a stated error history invites calibrated reliance (Akintola et al., 2025).

The theory of planned behavior explains the final step from belief to action: intention depends on attitude toward the behavior, perceived social norms, and perceived behavioral control, and action follows intention more reliably when control is real (Ajzen, 1991). A dashboard can raise perceived control by converting an abstract goal into a concrete next step, such as scheduling a transfer or adjusting a withholding form, and by reducing the friction between insight and execution. It can lower

perceived control if it presents problems without feasible responses, which produces avoidance; the so-called ostrich effect, in which people check accounts less often when markets fall or balances are low, is one expression of this. The planned behavior lens therefore predicts that prompts linked to executable actions outperform warnings alone, and that the benefit is concentrated among households with enough slack to act (Akhigbe et al., 2024; Eyetsemitan et al., 2023a).

3 Review Protocol

The study is an integrative review. It combines empirical studies, systematic reviews, regulatory and standards documents, and conceptual work in order to build theory where no single body of evidence is sufficient, and it reports its protocol so that the basis for each inference can be inspected. The reporting follows the elements of the PRISMA 2020 statement that apply to a theory-building review, namely a stated search strategy, eligibility criteria, appraisal approach and synthesis method, without claiming the quantitative pooling of a meta-analysis (Page, 2021; Dybå et al., 2008). The propositions derived from the synthesis are untested and are presented as such.

3.1 Sources, Search Strings and Time Window

Searches covered Scopus, Web of Science Core Collection, the ACM Digital Library, IEEE Xplore, EconLit and SSRN, together with reports from financial regulators, central banks, tax authorities and the OECD, and by backward and forward citation chasing from included studies. The window ran from January 2005 to December 2025; the start date captures the period in which online account aggregation and personal finance management applications became widely available, and the end date is the publication year of this article. The core search string combined three concept blocks with Boolean AND: a household block (household* OR consumer* OR "personal finance" OR taxpayer* OR famil*), a display block (dashboard* OR visuali* OR "account aggregat*" OR "open banking" OR "personal financial management" OR "budgeting app*" OR "decision aid*"), and an outcome block (saving* OR debt OR overdraft OR "cash flow" OR budget* OR tax* OR withholding OR "financial well-being" OR "financial capability"). A second string replaced the display block with behavioral terms ("present bias" OR "mental accounting" OR nudge* OR default* OR reminder* OR "commitment device*") to capture interventions delivered through digital channels (Easterbrook et al., 2008; Borg et al., 2014).

3.2 Eligibility, Screening and Appraisal

Records were eligible if the unit of analysis was an individual or household making a personal financial or tax decision; if the intervention or object of study was a digital display, aggregator, alert, projection or visual decision aid, or a behavioral mechanism that such displays operationalize; and if at least one outcome concerned comprehension, a decision, a financial behavior or a harm. Records were excluded if they addressed only enterprise reporting with no household analogue, reported usability or adoption without any decision or behavior outcome, or were promotional material without a described method. Organizational dashboard studies were admitted as a secondary stream where they informed a transferable design property, such as KPI definition or anomaly alerting, and were coded as analogical evidence (Filani et al., 2022; Sakyi et al., 2022b). Screening proceeded from titles and abstracts to full texts, with a coded reason retained for each exclusion.

Appraisal used a design-specific checklist rather than a single score. Randomized and natural experiments were appraised for allocation concealment, attrition, outcome timing and whether the measured outcome was behavior or stated intention. Observational studies of application users were appraised for selection into use, since households that adopt budgeting tools differ from those that do not in conscientiousness, income stability and prior literacy. Laboratory studies of chart comprehension were appraised for task realism and incentive alignment. Industry reports were appraised for disclosure of method and for conflicts of interest where the author sells the product evaluated. Standards and regulatory texts were treated as specifications of expected controls, not as evidence that those controls improve outcomes (ISO/IEC/IEEE, 2018; Yusuf et al., 2018).

3.3 Synthesis and Adjudication Logic

Extraction recorded, for each included source, the population, setting, display or intervention, comparator, construct measured, outcome, timing, effect direction and stated limitations, and coded each finding to one or more of the six LEDGER constructs. Synthesis followed a mechanism, context and outcome logic: a finding was interpreted by asking which mechanism it implicated and under which household and data conditions. Where findings conflicted, adjudication weighed three factors in order: design strength (experimental over quasi-experimental over cross-sectional), outcome proximity (observed behavior over stated intention over satisfaction), and population match (households over students over organizational analysts). Null and adverse findings were retained rather than discounted, because a framework that explains only successes cannot specify its

own failure conditions (Runeson et al., 2009; Arumosoye et al., 2019). Table 1 summarizes the protocol elements and the safeguard attached to each, and Figure 1 shows the sequence from identification to the outputs used in Sections 4 and 5.

Table 1. Review protocol for household dashboard and decision-quality evidence

Protocol element	Specification in this review	Safeguard against bias
Question scope	Household cash-flow, debt, saving and tax decisions supported by digital displays, aggregators or projections	Organizational dashboard evidence admitted only as coded analogical evidence
Sources	Six bibliographic databases plus regulator, central bank, tax authority and OECD reports; citation chasing	Grey literature appraised for method disclosure and vendor conflict of interest
Time window	January 2005 to December 2025	Early aggregator studies flagged where screen scraping preceded token-based access
Inclusion	Household unit; digital display or behavioral mechanism it operationalizes; comprehension, decision, behavior or harm outcome	Usability-only and adoption-only studies excluded with coded reason
Appraisal	Design-specific checklist covering allocation, attrition, selection into use, task realism, outcome timing	Stated intention and satisfaction not treated as decision outcomes
Coding	Findings mapped to six LEDGER constructs with mechanism, context, outcome and harm fields	Two coders; disagreements logged and resolved by discussion
Adjudication	Conflicts weighed by design strength, outcome proximity and population match	Null and adverse findings retained and reported in Table 2

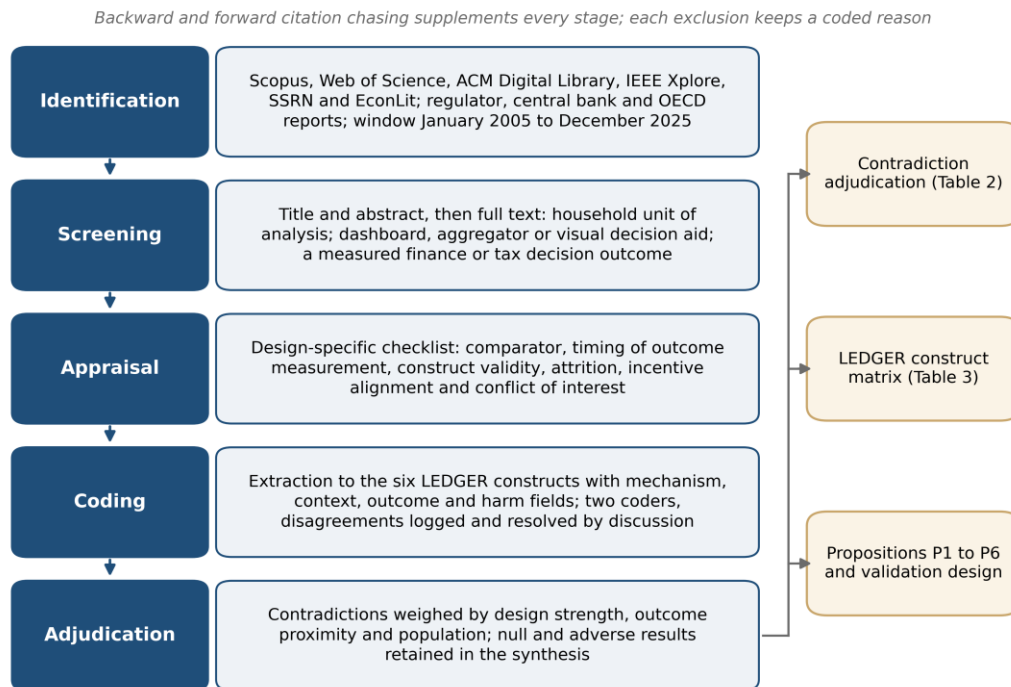


Figure 1. Review workflow from identification to adjudication and the three analytical outputs

Figure 1 makes explicit that the three outputs of the review, the contradiction adjudication, the construct matrix and the propositions, all derive from the adjudication stage rather than from any single included study. This sequencing matters because several influential claims about household finance applications rest on user surveys conducted by providers, which would dominate a synthesis that weighted sources by volume. Placing appraisal before coding ensures that each construct in the framework is anchored to the strongest available evidence for its mechanism, and that analogical evidence from

organizational business intelligence is labeled as such when it informs a design property (Dada et al., 2021a; Hevner et al., 2004).

4 Evidence Synthesis

The synthesis is organized by the decision pipeline a household dashboard implements: data are gathered and reconciled, encoded visually, framed relative to goals and categories, linked to action, qualified by uncertainty, and governed by consent. Each subsection reports what the evidence supports, where it is thin, and where it conflicts. A recurring observation is that the strongest evidence concerns individual mechanisms studied in isolation, such as reminders, defaults or chart types, whereas the weakest concerns integrated dashboards evaluated against household outcomes, which is the level at which products are sold (Miller et al., 2015; Tonoyan et al., 2024b). The subsections therefore distinguish, wherever the evidence allows, between findings about isolated mechanisms and findings about whole products, and between proximate outcomes such as comprehension and distal outcomes such as fees, penalties and savings balances (Sanni et al., 2022).

4.1 Aggregation, Categorization and Reconciliation

Every downstream display inherits the errors of aggregation. Consumer aggregators historically relied on screen scraping, in which the consumer supplied online banking credentials and the aggregator logged in on their behalf; open banking regimes have progressively replaced this with token-based application programming interfaces, under the revised Payment Services Directive in the European Union, the Open Banking implementation in the United Kingdom, and the personal financial data rights rule adopted in the United States in 2024. Token-based access improves security and consent granularity, but coverage remains uneven: pension accounts, employer equity plans, cash, informal loans and tax authority balances are often outside scope, so a net-worth figure can be precise and still materially incomplete (Mbonu et al., 2021a; World Bank, 2014).

Transaction categorization introduces a second error layer. Rule-based and machine-learned classifiers assign merchants to spending categories, but transfers between a household's own accounts, peer-to-peer payments, refunds and split purchases are routinely misclassified, and a transfer counted twice as spending distorts budgets and savings rates. Reconciliation, the matching of flows across accounts so that internal transfers net to zero, is a standard control in organizational finance and closes this gap, yet it is rarely described in consumer products (Dogbatsey et al., 2019; Oyeleye et al., 2025). Anomaly detection on household transaction series can flag duplicated charges, unexpected subscription renewals or fraudulent activity, but its false alarm rate must be low enough that users do not habituate, a problem well documented in industrial alarm management where excessive nuisance alarms degrade operator response (Atakpa et al., 2023; Ekelemu, 2025a).

The evidence therefore supports a design requirement rather than a feature: each displayed aggregate should be traceable to the transactions and accounts that produce it, with coverage disclosed. Traceability between requirements, data and outputs is a mature discipline in systems engineering, where bidirectional links let an analyst move from a reported figure to its sources and back (Gotel et al., 1994; Ramesh et al., 2001; Cleland-Huang et al., 2012). Transposed to household dashboards, traceability means that a user who doubts a savings rate can see which flows were counted, and that the provider can test the calculation engine against known ledgers. Automated recovery of such links is imperfect even in software repositories, which suggests that consumer systems should record lineage at ingestion rather than infer it afterward (De Lucia et al., 2007; Aliliele et al., 2024a).

4.2 Visual Encodings for Cash Flow, Debt and Savings

Household decisions involve three recurring quantitative questions: how the balance will evolve over the coming weeks, how the cost of debt compares across options, and how saving accumulates toward a target. Each maps onto a different encoding problem. Cash-flow trajectory is a time series, best shown as a line or step chart of projected balance with scheduled obligations marked on the time axis; stacked area charts of inflows and outflows are common but make the net position hard to read. Debt comparison is a ranking problem, best served by sorted bars of effective annual cost or total interest to payoff, whereas the common practice of listing balances with minimum payments emphasizes the least decision-relevant quantity. Savings progress is a part-to-whole problem where a single bar against a target outperforms circular progress rings for magnitude judgments (Ogbole et al., 2021; Filani et al., 2023).

Debt displays deserve particular attention because the evidence on debt literacy is discouraging. Many consumers cannot compute how long it will take to repay a balance at a given payment, and many underestimate the effect of compounding, patterns that are associated with higher fees and more expensive borrowing (Lusardi et al., 2015). Disclosure reforms that show the time and total cost of paying only the minimum have shifted some repayment behavior, but the shifts are modest and

concentrated among borrowers who were already paying near the minimum. A dashboard can go further by simulating alternative payments interactively, showing the months saved and interest avoided for each additional unit paid, and by ranking debts under both the highest-rate-first rule, which minimizes cost, and the smallest-balance-first rule, which some households find easier to sustain (Brown et al., 2016; Filani et al., 2023). Showing both, with the cost difference stated, respects the household's own trade-off between cost and motivation rather than imposing one.

The evidence on display density is mixed and is taken up in Section 4.6. Organizational studies of key performance indicator dashboards find that executives benefit from a small number of well-defined indicators with thresholds and owners, and that proliferating indicators dilutes attention (Sakyi et al., 2022b; Sanni et al., 2021b). Conceptual work on performance dashboards in operations management likewise ties their value to leadership visibility over a few decision-relevant measures, rather than to the volume of data on display (Ezeako et al., 2024). Household evidence points in the same direction for novice users but not for engaged ones, who explore detail and value drill-down. A layered design, with a small overview answering what needs attention this week and optional depth beneath, is the design most consistent with both findings, although it has rarely been tested head to head against flat alternatives (Tonoyan et al., 2023; Akanbi et al., 2024).

4.3 Behavioral Design: Mental Accounts, Commitment and Nudges

Digital channels have made several behavioral interventions cheap to deliver. Reminders timed before a payment due date reduce late fees; defaults that enroll a household in automatic saving raise participation far more than equivalent information; commitment devices that restrict withdrawals attract a subset of users who recognize their own present bias; and labeled sub-accounts for goals increase balances held for those goals. The consistent lesson is that interventions acting at the moment of choice, or that remove the need for a future choice, outperform interventions that deliver information and rely on the user to act later (Fobellah, 2025a; Xiao et al., 2016). This is exactly the prediction of the quasi-hyperbolic model in Section 2.1.

Financial education provides a useful contrast. Meta-analytic evidence indicates that education and counseling interventions improve some behaviors, notably saving and record keeping, while effects on other outcomes such as loan default are weaker and less consistent, and effects decay over time (Miller et al., 2015). State-mandated financial education for young people has been associated with improved credit behavior and lower delinquency, suggesting that education delivered at scale and close to the first borrowing decisions can matter (Brown et al., 2016). A field experiment with microentrepreneurs found that simple rules of thumb, such as keeping business and household money in separate drawers, improved financial practices more than standard accounting training (Drexler et al., 2014). Together these results imply that a dashboard's value may lie less in teaching and more in embedding good rules into the default structure of the display, for example by automatically separating a tax reserve from spendable balance for a self-employed user (Anichukwueze et al., 2022; Amarahobu et al., 2025a).

Behavioral design has documented limits. Nudges that work in one population may not transfer, effects measured over weeks may fade over years, and reminders sent too often produce habituation. There is also an ethical boundary: a dashboard operated by a lender or a product seller faces incentives to nudge toward its own products, and the same salience techniques that promote saving can promote spending or credit use. Consumer complaint data are a useful early signal of such misuse, in the way that complaint streams have been used to detect systemic risk in regulated service industries (Adeyelu, 2018; Fobellah, 2025c). The framework therefore treats guided action as a construct whose effect sign is an empirical question, and pairs it with harm monitoring in the validation design (Obriki et al., 2025).

4.4 Tax Views: Withholding, Estimated Payments and Readiness

Tax management is the part of household finance where dashboards are least studied and potentially most useful. In the United States, employees adjust withholding through a certificate filed with the employer, while self-employed and investment income generally requires quarterly estimated payments; an underpayment penalty applies unless payments meet safe-harbor tests, which for most filers require paying at least 90 percent of the current year's tax or 100 percent of the prior year's tax, rising to 110 percent for higher incomes. Other jurisdictions differ in mechanics, but the decision structure is similar: the household must forecast an annual liability from partial information and decide how much to set aside or prepay (Bola-Sadipe et al., 2022; Alm, 2019).

A tax view can support this decision in three ways. First, it can project year-end liability from year-to-date income and withholding, flagging when a second job, a bonus, or a capital gain has moved the household off its expected path. Second, it can translate the projection into the safe-harbor test, which is the decision-relevant quantity: the household needs to know not only the estimated liability but whether its current payments protect it from penalty. Third, it can schedule the cash required, linking tax payments to the cash-flow projection so that a quarterly payment does not collide with rent. The worked scenario

in Section 5.3 illustrates the third function. None of these functions eliminates uncertainty, since income may change and credits may phase in or out, so each projection should carry an interval and the assumptions behind it (Wedraogo et al., 2024; Komi, 2023).

Fiscal psychology adds a caution. Because many households value refunds as forced saving, a withholding optimizer that minimizes overpayment may reduce welfare as the household perceives it, and may be overridden. A more defensible design shows the refund the household is on track to receive, the cost of the implicit interest-free loan to the government, and an alternative automatic saving arrangement that would preserve the commitment benefit, leaving the choice explicit (Kirchler, 2007). Tax rules are also complex and change annually, so the calculation engine is a regression-testing problem: every rule change must be verified against a suite of known household cases, and coverage of edge cases such as phase-out thresholds matters more than coverage of common ones (Ammann et al., 2016; Myers et al., 2011; Kaner et al., 1999).

4.5 Privacy, Consent and Data Rights

Household dashboards concentrate the most revealing data a household produces. Transaction streams disclose health providers, religious donations, political contributions, location patterns and relationship changes, and tax records add income, dependants and filing status. Privacy harm can arise from breach, from secondary use by the provider or its partners, or from inference, as when spending patterns are used to predict creditworthiness or vulnerability. Privacy-by-design principles, role-based access control, data minimization and purpose limitation are well specified for enterprise systems handling personal data and transfer directly to consumer aggregators (Badmus et al., 2024; Ladapo et al., 2023b; Mbonu et al., 2018).

The evidence on how privacy concern affects dashboard use is instructive. Studies of digital services in health and finance report that trust, transparency about data use, and perceived control are stronger predictors of continued use than abstract privacy attitudes, and that users will share sensitive data when the benefit is concrete and the terms are understandable (Akintola et al., 2025; Kumuyi et al., 2024b). Consent that is granular, revocable and visible, for instance a panel listing which institutions are connected, which third parties receive data and when access expires, supports this control. Data protection impact assessment and data sensitivity classification provide the governance scaffolding, and lakehouse architectures with data loss prevention controls show how sensitive fields can be segregated from analytics workloads (Mbonu et al., 2022a; Aliliele et al., 2025b).

Security threats are not hypothetical. Credential sharing under screen scraping exposed households to account takeover; token architectures reduce but do not remove this risk, and the addition of conversational assistants that can act on a user's behalf introduces new vectors such as prompt injection and tool misuse (Nwakamma et al., 2024b). Zero-trust principles, continuous control monitoring and quantified risk reporting, developed for critical infrastructure, offer a model for how a provider should evidence its controls to regulators and users (Ahmed et al., 2021; Adebayo et al., 2025; Adegbite et al., 2025). Device-level risks also matter because household dashboards are consumed on personal phones shared among family members, a pattern analogous to bring-your-own-device risks documented in institutional settings (Onche et al., 2023; Jimoh et al., 2024). Where learning from multiple households is needed, federated approaches can keep raw transactions on the device or with the data holder (Sanni et al., 2023).

4.6 Adjudicating Conflicting Findings

Five contradictions recur in the evidence, and each bears directly on design. Table 2 states each question, the evidence on each side, and the adjudication reached under the rules in Section 3.3. The first concerns display density. Organizational KPI research and cognitive load theory favor few indicators, while engagement data from personal finance applications show that active users explore detailed views. The adjudication is that density should be layered and conditional on literacy, because the engagement evidence is observational and dominated by self-selected, higher-literacy users (Sakyi et al., 2022b; Tonoyan et al., 2024b).

The second contradiction concerns whether dashboards improve outcomes at all. Cross-sectional studies associate aggregator use with lower fees and higher saving, but the households that adopt such tools differ systematically from those that do not, so the association cannot be read causally. Experimental evidence on specific features such as alerts and reminders is more favorable than evidence on the full dashboard, which suggests that effects are concentrated in action-linked features. Cross-sectional risk indices in other fields face the same limitation, where an association between a composite marker and estimated risk does not establish that changing the marker changes the risk (Amadi et al., 2025; Easterbrook et al., 2008). The third contradiction concerns rules versus information: the rules-of-thumb evidence favors simple embedded heuristics for low-

literacy users, while financial capability research favors building understanding (Drexler et al., 2014; Xiao et al., 2016). The adjudication is that both can be served by defaults that embed a rule with an explanation available on demand.

The fourth contradiction concerns uncertainty display. Showing intervals should support calibrated reliance, but some users read intervals as noise and fall back on the central estimate, or interpret wide intervals as a sign of low quality and disengage. The weight of experimental evidence from forecast communication favors showing uncertainty in formats that express frequency, such as counts of simulated outcomes, over abstract intervals, particularly for lower-numeracy users (Wedraogo et al., 2024; Ilodigwe et al., 2024a). The fifth contradiction concerns fairness: forecasting and categorization models that are accurate on average can be systematically less accurate for households with irregular income, which are precisely those most exposed to shortfalls, so aggregate accuracy metrics can mask inequitable error (Komi, 2023; Uchechi et al., 2022).

Table 2. Conflicting findings on household dashboard design and their adjudication

Question	Evidence favoring	Evidence against or qualifying	Adjudication used in LEDGER
Do more indicators help?	Engaged users explore detail; drill-down valued in usage logs	KPI and cognitive load research: few indicators with thresholds improve attention	Layered overview with optional depth; density conditioned on literacy
Do dashboards improve outcomes?	Users show lower fees and higher saving in cross-sectional data	Selection into use; feature-level experiments stronger than whole-product evidence	Effects attributed to action-linked features; whole-product claims need experiments
Rules or understanding?	Rules of thumb outperform accounting training for low-literacy users	Capability gains require understanding that transfers to new decisions	Embed rules as defaults with explanation on demand
Should uncertainty be shown?	Intervals support calibrated reliance on forecasts	Some users ignore or distrust intervals; wide bands reduce engagement	Frequency formats and threshold probabilities rather than bare intervals
Is accuracy enough?	Aggregate forecast and categorization accuracy is high	Error concentrated among irregular-income households	Report error by income-volatility stratum; set per-stratum thresholds

The adjudications in Table 2 share a pattern. In each case the apparent conflict dissolves when the population and the outcome are specified: evidence that favors a feature typically comes from engaged, higher-literacy users and proximate outcomes such as comprehension, while evidence that qualifies it comes from harder-to-reach households and distal outcomes such as fees or penalties. This pattern is consistent with the broader finding that interventions which appear effective in average terms can widen gaps when uptake is unequal, a concern raised in research on diversity and heterogeneous team performance and in household economic stability studies (Amayo et al., 2023c; Jones et al., 2021). It motivates the moderators in the framework that follows.

5 The LEDGER Framework

The LEDGER framework translates the synthesis into six design constructs, three mediating states and three household outcomes, with moderators and a harm path. Its name is an acronym of the constructs: linked, reconciled data (L); encoding fit and literacy (E); decision anchoring through mental accounts and planning horizon (D); guided action and nudges (G); exposed uncertainty (E); and rights over data (R). The constructs are not a maturity sequence and are not additive; the framework predicts interactions, for example that guided action without exposed uncertainty raises over-reliance, and that encoding fit matters more for low-literacy users. Figure 2 shows the structure and Table 3 operationalizes each construct (Hevner et al., 2004; Arumosoye et al., 2021).

5.1 Constructs and Measurement

Each construct is defined by what it changes in the household's information environment and is measured by an indicator that can be computed from system logs or an experimental task, rather than by user perception alone. Linked data is measured by coverage (the share of the household's balance sheet and cash flows captured) and by reconciliation error (the share of internal transfers misclassified as income or spending). Encoding fit is measured by task accuracy on standard chart-reading items using the dashboard's own encodings, stratified by the user's visualization literacy. Decision anchoring is measured by whether the display places each decision relative to a stated goal and horizon, and by the rate at which users make cross-account moves that reduce total cost. Guided action is measured by the share of prompts that link to an executable step and by prompt-to-action conversion within a defined window. Exposed uncertainty is measured by calibration of the user's probability

judgments against the model's outcomes. Rights over data are measured by consent granularity, the time to revoke access, and the proportion of users who can correctly state which parties receive their data (Netemeyer et al., 2018; Sakyi et al., 2022b).

Table 3. LEDGER constructs, indicators, failure signals and design responses

Construct	Operational indicator	Failure signal	Design response
L: Linked, reconciled data	Balance-sheet coverage; internal-transfer misclassification rate; lineage available per aggregate	Net worth or savings rate shifts when a transfer is recategorized	Disclose coverage; reconcile before display; show source transactions on demand
E: Encoding fit and literacy	Chart-task accuracy with the product's encodings, by literacy tercile	Accuracy gap between literacy terciles exceeds a set tolerance	Replace angle and area encodings with position; add plain-language readouts
D: Decision anchoring	Share of views anchored to a goal and horizon; rate of cost-reducing cross-account moves	Idle savings coexist with revolving high-cost debt	Consolidated trade-off view showing the net cost of the current allocation
G: Guided action and nudges	Share of prompts linked to an executable step; prompt-to-action conversion within seven days	High prompt volume with falling conversion (habituation)	Throttle prompts; tie each to one action; audit prompts for seller conflict
E: Exposed uncertainty	Brier score of users' shortfall probability judgments; forecast interval coverage	Users treat central estimates as guarantees; intervals miscalibrated	Frequency-format uncertainty; state assumptions; publish error history
R: Rights over data	Consent granularity; median time to revoke; accuracy of users' data-sharing beliefs	Users cannot name recipients of their data; revocation buried	Consent dashboard with expiry; minimization; per-purpose toggles

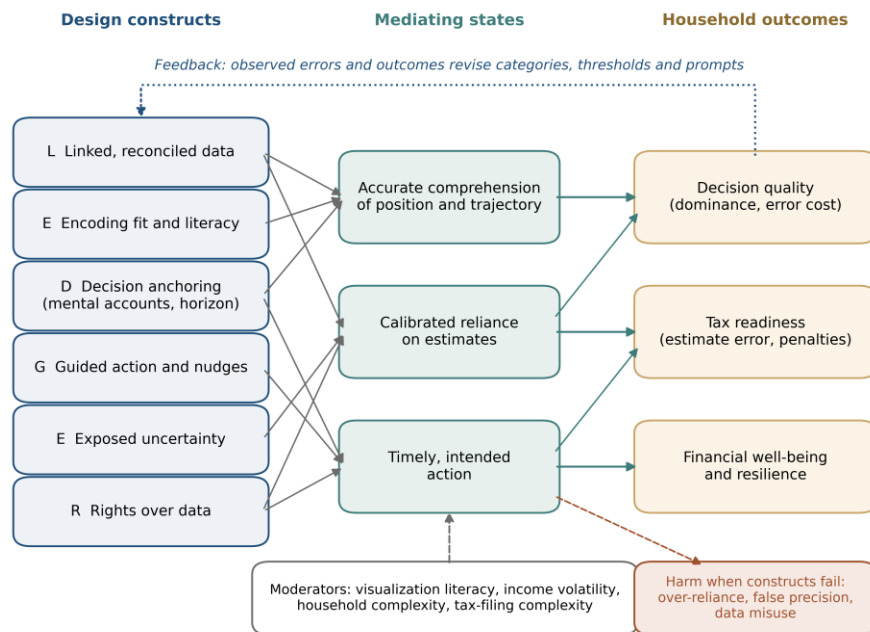


Figure 2. The LEDGER framework linking six design constructs to mediating states and household outcomes, with moderators, harm path and feedback

Figure 2 separates the mediating states from the outcomes because the evidence in Section 4 shows that dashboards often move the former without moving the latter. Accurate comprehension of position and trajectory is necessary for good decisions but not sufficient; calibrated reliance on estimates is necessary for tax readiness because tax projections are uncertain; and timely, intended action is the route through which information changes financial well-being. The feedback path indicates that observed errors, such as recategorized transactions or ignored prompts, should revise the categories, thresholds and prompts themselves, which is how learning organizations use near-miss data to improve controls (Arumosoyle et al., 2019; Obriki et al.,

2019). The harm path makes explicit that each construct, if poorly implemented, has a failure mode that the evaluation must detect.

5.2 Formal Decision-Quality Metrics

Decision quality should be measured against a normative benchmark where one exists and against the household's own stated goals where it does not. Four metrics follow from the synthesis. The first is the dominance error rate. In a task where one option is better on every dimension the household cares about, such as repaying a 24 percent card balance rather than holding the same sum in a 1 percent savings account while an emergency buffer is already funded, choosing the dominated option is an unambiguous error. Across a set of N tasks with dominance structure, the rate is $DER = (1/N) \sum_i 1[\text{choice}_i \text{ is dominated}]$. Dominance tasks avoid imposing a particular utility function while still identifying clear mistakes (Drexler et al., 2014; Li et al., 2007).

The second metric is the expected cost of decision error. For tasks where the cost-minimizing choice is known, such as selecting which debt to prepay, the cost of error is $ECE = (1/N) \sum_i [C(a_i) - C(a_i^*)]$, where $C(a_i)$ is the total cost of the chosen action and $C(a_i^*)$ is the cost of the optimal action, both computed over a stated horizon. ECE weights errors by their financial consequence, so a dashboard that eliminates small errors but not large ones will score poorly. The third metric is the Brier score of probabilistic judgments. Users are asked for the probability that their balance will fall below a buffer within a stated period; with judged probability p_i and realized outcome o_i in $\{0,1\}$ across simulated or observed episodes, $BS = (1/N) \sum_i (p_i - o_i)^2$, and lower scores indicate better calibrated beliefs. The Brier score can be decomposed into reliability, resolution and uncertainty components, which distinguishes users who are poorly calibrated from those who simply cannot discriminate risky from safe periods (Wedraogo et al., 2024; Atakpa et al., 2023).

The fourth metric is tax estimate error relative to the safe harbor. Let T be the realized annual liability, T_{hat} the household's estimate at a given date, and P the payments made or scheduled. The absolute percentage error $|T_{\text{hat}} - T| / T$ measures estimation accuracy, but the decision-relevant indicator is whether P meets the applicable safe-harbor threshold, an indicator $SH = 1[P \geq \min(0.9 T, k T_{\text{prior}})]$, where k is 1.0 or 1.1 depending on prior-year income and T_{prior} is the prior year's liability. A tax view succeeds when it raises the share of households meeting SH without inducing large overpayments, so both SH and the overpayment ratio $(P - T) / T$ should be reported (Bola-Sadipe et al., 2022; Catal et al., 2013). These four metrics complement, rather than replace, validated self-report scales of perceived financial well-being, which capture the household's own evaluation of stress and future security (Netemeyer et al., 2018).

5.3 Worked Scenario: A Variable-Income Household

Consider a household with a biweekly salary, additional gig income that varies from week to week, occasional lumpy expenses such as car repairs, and quarterly estimated tax payments on the gig income. Figure 3 shows a simulated projection of this household's combined checking and savings balance over a tax year, constructed from 4,000 Monte Carlo paths with normally distributed weekly shocks and randomly timed lumpy expenses. The figure is illustrative and does not describe a real household. The median path rises between tax payments and drops at each quarterly payment. The 90 percent interval widens over time as uncertainty accumulates, and the probability that the balance falls below a 1,500-dollar buffer first exceeds 10 percent in the week of the second estimated payment (Filani et al., 2023; Aminu-Ibrahim et al., 2025b).

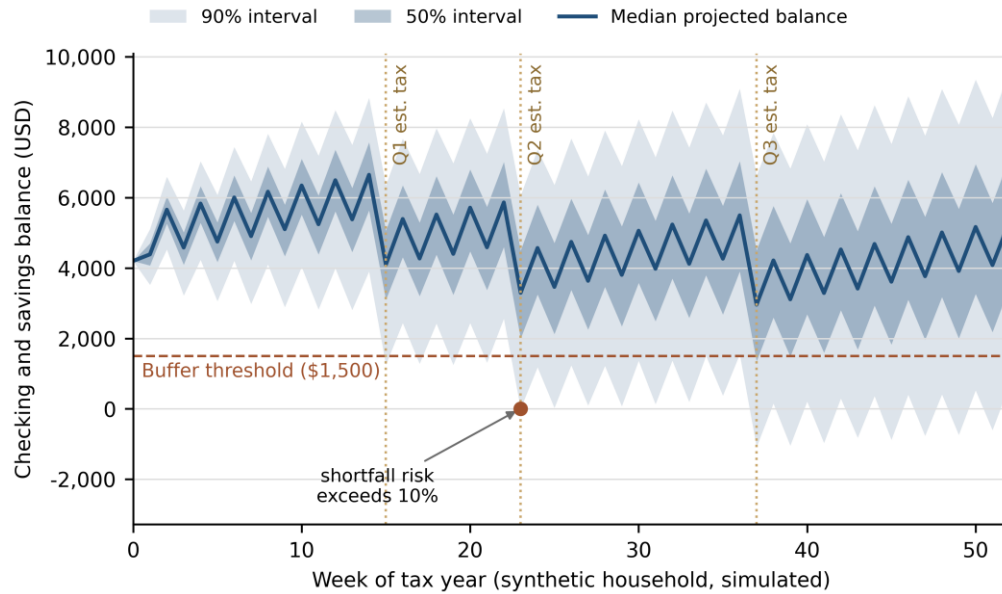


Figure 3. Illustrative projected balance for a simulated variable-income household, with 50 and 90 percent intervals, quarterly estimated tax payments and a buffer threshold

The scenario clarifies what each LEDGER construct contributes. Without linked data the gig income would be missing from the projection, and without reconciliation a transfer to savings would appear as spending. With a single-line forecast the household would see a comfortable median and no warning; the exposed-uncertainty construct converts the interval into a decision-relevant quantity, namely the week in which shortfall risk becomes material. Decision anchoring places the tax payment inside the household's cash-flow view rather than in a separate tax tile, so the collision between the payment and a low-balance week is visible. Guided action then offers concrete responses, such as moving part of the third-quarter payment earlier while the balance is high, raising withholding on the salary to reduce the estimated payments, or building the buffer from the next two paychecks, each with its effect on shortfall probability shown. Rights over data determine whether the household was willing to connect the gig platform in the first place (Aminu-Ibrahim et al., 2025a; Ekelemu et al., 2025b).

The scenario also shows a trade-off that aggregate metrics conceal. A forecasting model trained mainly on salaried households will understate variance for this household, narrowing the interval and delaying the warning. Its average error across all users may still look acceptable, which is the fairness concern raised in Section 4.6. Reporting calibration separately for strata of income volatility is therefore part of the framework, not an optional refinement (Komi, 2023; Aminu-Ibrahim et al., 2025c).

5.4 Propositions and Tests

Six propositions follow. Each names the comparison that would test it and the result that would count against it. P1 (encoding fit): among households in the lowest visualization-literacy tercile, dashboards using position encodings for spending, debt and savings will produce lower dominance error rates than dashboards using angle, area and color encodings for the same data; the difference will be smaller in the highest tercile. The proposition is refuted if the encoding effect does not differ by literacy or favors the compact encodings. P2 (exposed uncertainty): frequency-format uncertainty displays will produce lower Brier scores for shortfall judgments than single-line forecasts, without increasing time on task by more than a pre-specified margin. It is refuted if Brier scores do not improve or if users disengage. P3 (guided action): prompts linked to an executable action will produce higher seven-day action rates and lower expected cost of error than equivalent warnings without an action link (Ajzen, 1991; Eze et al., 2024b).

P4 (interaction): the benefit of guided action on expected cost of error will be smaller, and over-reliance on erroneous recommendations will be larger, when uncertainty is not exposed; the test is a factorial contrast in which some recommendations deliberately contain planted errors and acceptance of erroneous recommendations is measured. P5 (tax readiness): in field use, households with a tax view that expresses projections relative to the safe-harbor threshold will meet the threshold more often than comparable households with a liability estimate alone, without a larger overpayment ratio. It is refuted if safe-harbor attainment does not rise or rises only through large overpayment. P6 (rights over data): households given

a granular, revocable consent panel will connect more accounts and revoke less often than households given an all-or-nothing consent flow, because perceived control raises willingness to share when the benefit is concrete. It is refuted if granular consent reduces coverage or leaves revocation unchanged (Kumuyi et al., 2024b; Akhigbe et al., 2025). Each proposition permits a null or adverse result, and the validation design below is built to detect both (Ilodigwe et al., 2024b).

6 Validation Design

The propositions call for two kinds of evidence: controlled comparisons of display variants on tasks where the normative answer is known, and field evidence that effects survive real accounts, real stakes and the passage of time. The design in Figure 4 combines the two in sequence, so that only variants that pass the laboratory stage proceed to the field, which reduces the cost and risk of field deployment. The approach follows design science practice in treating the framework as an artifact to be evaluated against explicit criteria, and follows longitudinal evaluation practice in measuring whether early gains persist (Wieringa, 2014; Onwuzulike et al., 2024).

6.1 Stage 1: Randomized Decision Experiment

Stage 1 recruits an online panel stratified by visualization literacy and by self-reported income volatility, and assigns participants at random to one of four display arms operating on identical synthetic household ledgers: a tabular statement (control), the full LEDGER dashboard, the dashboard without uncertainty bands, and the dashboard without action-linked prompts. Participants complete a battery of tasks with known normative answers, including debt prioritization, buffer adequacy, a cross-account allocation with a dominated option, and a safe-harbor judgment for a projected tax liability. Outcomes are the dominance error rate, expected cost of error, Brier score for shortfall probability judgments, safe-harbor choice accuracy and time on task. Incentives tied to task accuracy reduce the gap between stated and revealed choice. Some recommendations contain planted errors to measure over-reliance, as required by P4. The planned contrasts are the full dashboard against control, against the no-uncertainty arm and against the no-prompt arm, with literacy as a pre-registered moderator (Oluwadele et al., 2025; Daka et al., 2014).

Several design details protect the inferences drawn from Stage 1. The synthetic ledgers are generated from parameterized templates so that each participant sees a household of the same structure with different surface details, which prevents memorization and allows the normative answer to be computed exactly for every task. Task order is counterbalanced, and attention checks are embedded in the ledgers rather than presented as separate items. Power is planned for the literacy interaction in P1, which is the smallest effect of interest, rather than for the main effects, because an interaction powered only for main effects will usually be missed. Visualization literacy is measured before assignment with a short standardized chart-reading instrument, and financial literacy with the established items on interest, inflation and diversification, so that the two moderators can be separated (Atkinson et al., 2012; Obogo et al., 2019).

6.2 Stage 2: Stepped-Wedge Field Evaluation

Stage 2 deploys the surviving design to consenting households who connect their own accounts. A stepped-wedge design rolls out the LEDGER dashboard to successive waves of households at randomized times, so that every wave eventually receives it and each wave serves as a control for the others before crossover. This design is ethically attractive when the intervention is expected to help and all participants will eventually receive it, and it allows secular trends such as tax-season effects to be separated from the intervention effect. Behavioral outcomes are drawn from account data: overdraft and late-fee incidence, the number of days on which the balance exceeds the household's buffer, and, after filing, whether an underpayment penalty was incurred. Harm outcomes are monitored in parallel: consent revocations, reported data errors, complaints, and prompt acceptance for products sold by the provider. Process tracing links each prompt exposure to subsequent actions so that effects can be attributed to mechanisms rather than inferred from outcomes alone (Arumosoye et al., 2023; Obogo et al., 2019b; Sanni et al., 2025).

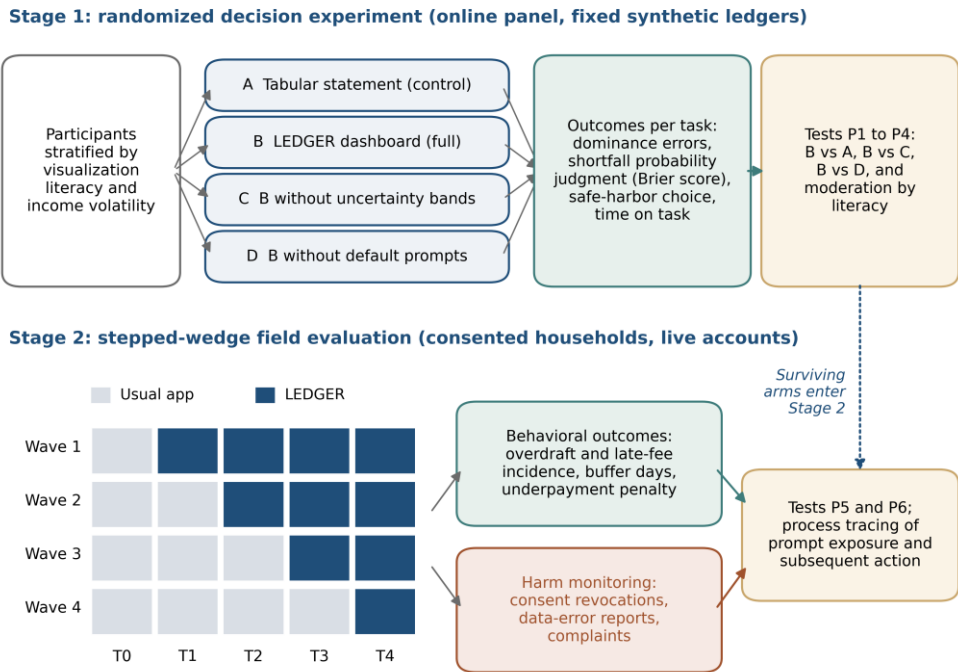


Figure 4. Two-stage validation design: a four-arm randomized decision experiment followed by a stepped-wedge field evaluation with parallel harm monitoring

Figure 4 shows that Stage 1 tests P1 to P4, where the normative answer is known and the display can be manipulated cleanly, and Stage 2 tests P5 and P6, which require real tax outcomes and real consent choices. The dotted link between the stages expresses a stopping rule: an arm that fails to reduce errors in Stage 1, or that increases acceptance of erroneous recommendations, does not proceed to households' real accounts. Continuous monitoring of the harm indicators during Stage 2, with predefined thresholds that pause rollout, applies the logic of leading indicators used in safety management to consumer financial risk (Obriki et al., 2025; Obogo et al., 2019a).

6.3 Threats to Validity

Each stage faces threats that a larger sample would not repair, and Table 4 lists those specific to household dashboard research together with the mitigation adopted. Construct threats arise because comprehension and decision quality can be conflated with familiarity: users of a given application will perform better on its encodings regardless of their merit. Internal threats in the field stage include tax-season seasonality and concurrent changes in income. External threats arise because online panels over-represent digitally confident users, and because tax rules differ across jurisdictions. Ethical threats include the possibility that the evaluation itself exposes data or nudges participants toward costly choices (Runeson et al., 2009; Mairiza et al., 2010; Fadayomi et al., 2019).

Table 4. Threats to validity in evaluating household financial dashboards and mitigations

Validity type	Threat specific to household dashboards	Mitigation in the validation design
Construct	Prior familiarity with a product's encodings inflates comprehension scores; satisfaction mistaken for decision quality	Novel synthetic ledgers; dominance and cost-of-error tasks with known answers; satisfaction reported separately
Internal	Tax-season and paycheck-cycle effects coincide with rollout; income shocks confound behavior change	Stepped-wedge randomization of timing; calendar fixed effects; income change recorded as covariate
External	Online panels over-represent high-literacy, stable-income users; safe-harbor rules are jurisdiction specific	Stratified recruitment by literacy and volatility; tax tasks parameterized by jurisdiction
Statistical conclusion	Rare outcomes such as underpayment penalties give low power; many outcomes invite selective	Pre-registered primary outcomes; pooled penalty

Validity type	Threat specific to household dashboards	Mitigation in the validation design
	reporting	and fee incidence; report all arms and nulls
Ethical	Field arm exposes sensitive data or steers households toward provider products	Data minimization and on-device processing; independent prompt audit; pause thresholds on harm indicators

The mitigations in Table 4 are not exhaustive, but they target the threats most likely to produce a misleading positive result. The use of synthetic ledgers in Stage 1 removes both privacy exposure and familiarity effects while allowing exact computation of the normative answer. The independent audit of prompts in Stage 2 addresses the concern that a provider's commercial incentives could shape the guided-action construct. Recording implementation fidelity, including whether households actually saw each prompt, is necessary for distinguishing a failed mechanism from a mechanism that never operated (Borah et al., 2022; Zhu et al., 1997; Ogbete et al., 2023).

7 Discussion

The review set out to identify which dashboard design properties improve household financial decisions, when they fail, and how decision quality should be measured. Its answer is that the value of a household dashboard lies in a small number of properties that act on the decision itself: reconciled data, encodings the user can decode, anchoring of choices to goals and horizons, prompts linked to action, uncertainty expressed in usable form, and control over data. The following subsections consider what this implies for theory and practice, the strongest counter-arguments, and the limits of the review (Hevner et al., 2004; Filani et al., 2022).

7.1 Theoretical Implications

The review has three implications for theory. First, it suggests that behavioral household finance and visual analytics should be treated as a single explanatory system for dashboard effects. Present bias predicts that information delivered away from the moment of choice will be discounted; graphical perception predicts that information delivered at the moment of choice will be misread if poorly encoded. Neither theory alone explains why a well-encoded monthly report fails, or why a timely prompt with a misleading chart succeeds only among expert users. Second, it extends fiscal psychology to the interface: the slippery slope framework's distinction between power and trust can be observed in how a tax projection frames the authority, and in whether households treat an estimated liability as a threat or a plan (Kirchler, 2007; Alm, 2019).

Third, the framework reframes financial capability as partly a property of the household's information environment rather than only of the individual. Capability research has measured knowledge and skill, and education research has tried to raise them, with modest and decaying effects (Miller et al., 2015; Xiao et al., 2016). If appropriately designed displays and defaults reduce decision errors among households whose knowledge is unchanged, then capability is co-produced by person and environment, and measures of literacy alone will understate the capability that a well-designed environment supports. This is analogous to how organizational performance depends on the design of reporting systems and control routines, not only on the skill of individual managers (Akhigbe et al., 2025; Falemi et al., 2024; Obogo et al., 2023).

7.2 Practical Implications

For providers, the most direct implication is to evaluate dashboards by decision quality and harm incidence rather than by engagement. Session counts and feature adoption are easy to measure and easy to raise, but they reward designs that attract attention rather than designs that improve decisions. A provider applying LEDGER would publish coverage and reconciliation error, test chart comprehension across literacy strata before release, attach each alert to a single executable action, show uncertainty for every forecast and tax projection, and offer a consent panel with expiry. Tax calculation engines should be subject to regression testing against a maintained suite of household cases whenever rules change, following established software testing processes (ISO/IEC/IEEE, 2018; Ammann et al., 2016; Eyetsemitan et al., 2023b).

For regulators and consumer protection bodies, the framework suggests supervisory questions. Does the provider disclose which accounts are outside coverage? Are prompts audited for conflicts with the provider's own products? Can a user revoke access in a single step? Are forecasting errors reported for irregular-income households separately? Where dashboards employ machine learning to categorize transactions or recommend actions, emerging artificial intelligence regulation that classifies creditworthiness assessment as high risk provides a template for documentation and human oversight obligations (Ominyi et al., 2024b; Ominyi et al., 2025a). For tax authorities, the framework suggests collaboration with software providers on year-

round projection tools, since reducing misestimation reduces both penalties and administrative cost, and because trust-building interactions support voluntary compliance (Kirchler, 2007; Uchechi et al., 2018).

7.3 Counter-Arguments and Boundary Conditions

Three objections follow naturally from the household setting. The first is that households with the greatest need, those with low and volatile incomes, have the least slack to act on any display, so dashboards will mostly benefit households that were already doing well. The evidence partly supports this, which is why income volatility is a moderator and why P5 and the field outcomes are stratified. It does not follow that dashboards are useless for constrained households: for them, a well-timed warning of a shortfall can prevent an overdraft fee that is large relative to income, so the marginal value of information may be highest where slack is lowest (World Bank, 2014; Lusardi et al., 2015).

The second counter-argument is that simpler interventions, such as an automatic transfer to savings or a text reminder before a due date, capture most of the benefit at a fraction of the cost and privacy exposure. This is a serious rival explanation, and the validation design includes a tabular control and a no-prompt arm partly to test it. If the full dashboard does not outperform a reminder-plus-default baseline on decision quality, the framework would recommend the simpler design, and the constructs would remain useful as a checklist for it (Drexler et al., 2014; Eze et al., 2022b). The third counter-argument is that the framework assumes a single decision-maker, whereas many households share finances across partners with different goals, literacy and access. Shared dashboards can reduce information loss between partners, much as structured handovers reduce information loss between work shifts, but they can also become instruments of financial control within a relationship, so access design must allow private as well as shared views (Ekelemu, 2025b; Ogbete et al., 2018). The boundary conditions of the framework follow from these points: it applies where households have digital access, where the relevant accounts can be connected, and where at least some decisions are within the household's control.

7.4 Limitations and Future Research

The review has limitations. It is integrative rather than meta-analytic, so it does not estimate effect sizes, and its adjudications reflect judgments about design strength and outcome proximity that other reviewers might weigh differently. The evidence base is concentrated in high-income countries with mature open banking or aggregation infrastructure, and tax evidence is dominated by a small number of jurisdictions, so transfer to settings with large informal sectors and cash economies is uncertain. Much of the evidence on integrated consumer dashboards comes from providers, whose incentives favor positive findings, and negative results are likely underreported. The propositions are untested, and the metrics, although grounded in established methods, have not been validated together in a household dashboard study (Page, 2021; Dybå et al., 2008; Atta et al., 2021).

Four directions for future research follow. First, the Stage 1 experiment should be run across jurisdictions with different tax prepayment regimes, to test whether the safe-harbor framing generalizes. Second, longitudinal field studies should track whether decision-quality gains persist beyond the first tax year, or decay as novelty fades, as financial education effects do. Third, research should examine households whose finances are partly informal, such as those relying on remittances, rotating savings groups or agricultural income, where aggregation coverage is inherently limited and where analytics has been proposed to support decisions in adjacent domains (Michael et al., 2025c; Jones et al., 2021). Fourth, as conversational assistants are embedded in financial applications, research should test whether natural-language explanations improve calibration relative to visual uncertainty displays or instead increase over-reliance, and how such assistants can be protected against manipulation (Nwakamma et al., 2024b; Ilodigwe et al., 2025a). Across all four, measuring organizational and individual learning from errors, rather than only outcome levels, would reveal whether dashboards help households improve their own decision rules over time (Arumosoye et al., 2024; Arumosoye et al., 2025; Obriki et al., 2020; Atta et al., 2022; Ogbete et al., 2022).

8 Conclusion

Household financial dashboards can improve decisions, but only through specific properties that act on the decision itself. This review has shown that reconciled data, perceptually efficient encodings matched to the user's visualization literacy, anchoring of choices to goals and horizons, prompts that lead directly to action, uncertainty expressed in usable form, and genuine control over personal data are the conditions under which aggregation and visualization translate into better cash-flow, debt, saving and tax decisions. The LEDGER framework organizes these conditions into constructs that can be manipulated and measured, and the proposed metrics, namely dominance errors, the expected cost of error, Brier-scored shortfall judgments

and safe-harbor attainment, give providers, regulators and researchers a common basis for comparing designs. The six propositions and the two-stage validation design specify how the framework can be confirmed or refuted. Judged by decision quality and harm incidence rather than by engagement, household dashboards can become a dependable part of how families plan, borrow, save and meet their tax obligations.

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

Competing interests: None declared.

Data availability: No new data were generated or analysed in this study; all sources are cited in the reference list. Figure 3 is a simulation described in the text.

References

- Abah, M. A., Okwah, M. N., Oladosu, M. A., Yohanna, N. R., Blessing, C. T., & Vanessa, O. I. (2025). The impact of lifestyle modifications on cardiovascular health: A review of diet, exercise, and stress reduction intervention. *Drug and Pharmaceutical Science Archives*, 5(3), 26-34. <https://doi.org/10.47587/DAP.2025.5301>
- Adebayo, A., Amadi, C., & Ijagbemi, A. O. (2025). From audit to continuous control: Implementation pathways, governance, and organizational readiness for AI-enabled compliance in critical infrastructure. *International Journal of Computer Science and Mathematical Theory*, 11(8), 88-134. <https://doi.org/10.56201/ijcsmt.vol.11.no8.2025.pg88.134>
- Adegbite, M. P., Adebayo, A., & Ahmed, M. O. (2025). A cyber risk quantification and governance architecture for critical infrastructure: From posture measurement to executive reporting. *International Journal of Computer Science and Mathematical Theory*, 11(12), 232-293. <https://doi.org/10.56201/ijcsmt.vol.11.no12.2025.pg232.293>
- Adelanwa, A., Basnet, A., & Anene, U. N. (2024). Advanced AI based decision support systems for healthcare operations and resource planning. *Gyanshauryam, International Scientific Refereed Research Journal*, 7(1), 244-276. <http://doi.org/10.32628/GISRRJ>
- Adeyelu, O. O. (2018). A predictive compliance monitoring framework for detecting systemic safety risks through aviation consumer complaint data. *Iconic Research and Engineering Journals*, 1(8), 250-276. <https://doi.org/10.64388/IREV118-1718478>
- Ahmed, M. O., Adegbite, M. P., & Adebayo, A. (2021). Zero trust architecture for operational technology in North American critical infrastructure: A framework for implementation and resilience optimization. *International Journal of Engineering and Modern Technology*, 7(1), 66-113. <https://doi.org/10.56201/ijemt.vol.7.no1.2021.pg66.113>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Akanbi, O., & Sunday, E. A. (2024). An integrated path-planning and slotting optimization model for AMR-enabled high-density warehousing. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3226-3243. <https://doi.org/10.62225/2583049X.2024.4.6.6181>
- Akhigbe, R., Falemi, A., & Akin-Oluyomi, O. T. (2024). An advanced framework for technology adoption that supports employee readiness in modern organizations. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(1), 1637-1659. <https://doi.org/10.62225/2583049X.2024.4.1.5344>
- Akhigbe, R., Falemi, A., & Akin-Oluyomi, O. T. (2025). An advanced governance model for improving transparency accountability and strategic execution in complex enterprises. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 1169-1188. <https://doi.org/10.62225/2583049X.2025.5.6.5341>

- Akintola, A. S., Fawehinmi, Y. O., Aiyenitaju, O., Chinemerem, B., & Emmanuella, O. (2025). Digital transformation in telehealth: A systematic review of user trust, privacy, and regulatory governance in AI-powered remote monitoring systems. *Journal of Scientific Research and Reports*, 31(11), 163-182. <https://doi.org/10.9734/jsrr/2025/v31i113658>
- Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2024a). A conceptual framework for enterprise data sensitivity classification and regulatory traceability mechanisms. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3103-3124. <https://doi.org/10.62225/2583049X.2024.4.6.5991>
- Aliliele, C., Mbonu, I. S., Uzoka, E., & Iwuanyanwu, U. (2025b). Advances in data lakehouse governance architectures for enterprise data loss prevention and compliance assurance. *Shodhshauryam, International Scientific Refereed Research Journal*, 8(4), 171-213. <https://doi.org/10.32628/SHISRRJ258474>
- Alm, J. (2019). What motivates tax compliance? *Journal of Economic Surveys*, 33(2), 353-388. <https://doi.org/10.1111/joes.12272>
- Amadi, C. E., Okorafor, U. C., Okorafor, C. I., Okwah, M. N., Akanbi, P. O., Okam, O. V., Udenze, C. I., Mbakwem, A. C., & Ajuluchukwu, J. N. (2025). Association between triglyceride-glucose index and estimated 10-year risk of cardiovascular disease in Nigerian non-diabetic hypertensives: A cross-sectional study. *PLOS Global Public Health*, 5(7), e0004760. <https://doi.org/10.1371/journal.pgph.0004760>
- Amarahobu, C., Ekelemu, O., Akano, O. A., & Adikwu, F. E. (2025a). Structured mentorship models in on-the-job technician development: A review of advancement and promotion coaching schemes, workplace learning theory and the leadership interface in continuous processing environments. *International Journal of Social Sciences and Management Research*, 11(8), 626-656. <https://doi.org/10.56201/ijssmr.vol.11no8.2025.pg626.656>
- Amayo, E. B., Owulade, O. A., & Isi, L. R. (2023c). The role of diversity and inclusion in enhancing project team performance and delivery in multinational environments. *Journal of Frontiers in Multidisciplinary Research*, 4(1), 48-58. <https://doi.org/10.54660/IJFMR.2023.4.1.48-58>
- Aminu-Ibrahim, A. Y., Ogbete, J. C., & Iwuanyanwu, O. C. (2025a). Cost control and financial accountability frameworks for national healthcare construction programs. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 1970-1990. <https://doi.org/10.62225/2583049X.2025.5.6.5700>
- Aminu-Ibrahim, A. Y., Ogbete, J. C., & Iwuanyanwu, O. C. (2025b). Infrastructure resilience planning for national diagnostic systems under public health stress conditions. *Gyanshauryam, International Scientific Refereed Research Journal*, 8(1), 340-381. <https://doi.org/10.32628/GISRRJ2582311>
- Aminu-Ibrahim, A. Y., Ogbete, J. C., & Iwuanyanwu, O. C. (2025c). Sustainable healthcare infrastructure performance metrics for long-term asset management and value creation. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 11(4), 566-601. <https://doi.org/10.32628/CSEIT251116277>
- Ammann, P., & Offutt, J. (2016). *Introduction to Software Testing* (2nd ed.). Cambridge University Press.
- Anichukwueze, C. C., Ominyi, M., Uzougbo, N. S., & Jones, B. C. (2022). Cross-functional legal education as risk mitigation: Empirical analysis of contract error reduction in engineering teams. *International Journal of Engineering and Modern Technology*, 8(5), 142-180. <https://doi.org/10.56201/ijemt.v8.no5.2022.pg142.180>
- Anioke, S. C., & Atima, M. E. (2023a). Public health governance models using process optimization and performance metrics for regulatory oversight. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2534-2548. <https://doi.org/10.62225/2583049X.2023.3.6.5491>

- Arumosoye, O. M., & Obriki, O. D. (2023). Conceptual model for emergency response readiness and capability in energy and process facilities. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(3), 897-917. <https://doi.org/10.32628/CSEIT25112791>
- Arumosoye, O. M., & Obriki, O. D. (2024). Conceptual model of safety leadership influence in large temporary project organizations. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3034-3047. <https://doi.org/10.62225/2583049X.2024.4.6.5895>
- Arumosoye, O. M., & Obriki, O. D. (2021). Organizational learning-based conceptual maturity model for continuous safety performance improvement. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(1), 970-980. <https://doi.org/10.54660/IJMRGE.2021.2.1.970-980>
- Arumosoye, O. M., Obriki, O. D., & Ozobu, C. O. (2025). Conceptual framework for environmental risk control and ESG performance through waste handling and operational discipline. *Gyanshauryam, International Scientific Refereed Research Journal*, 8(4), 177-198. <https://doi.org/10.32628/GISRRJ258367>
- Arumosoye, O. M., & Obriki, O. D. (2019). Systematic review of near-miss and hazard observation data utilization in industrial safety management. *Iconic Research and Engineering Journals*, 3(2), 981-999. <https://doi.org/10.64388/IREV3I2-1714417>
- Atakpa, M. I., & Fobellah, A. N. (2023). Anomaly detection in financial time-series data: A conceptual model for healthcare and banking applications. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(2), 967-997. <https://doi.org/10.32628/CSEIT2342441>
- Atkinson, A., & Messy, F. (2012). Measuring financial literacy: Results of the OECD/INFE pilot study. *OECD Working Papers on Finance, Insurance and Private Pensions*, 15. <https://doi.org/10.1787/5k9csfs90fr4-en>
- Atta, S., Adjaklo, E., Asomani, E., & Adenuga, O. M. (2022). Advances in welding inspection standards and quality assurance practices in large-scale power generation projects. *International Journal of Engineering and Modern Technology*, 8(5), 101-124. <https://doi.org/10.56201/ijemt.v8.no5.2022.pg101.124>
- Atta, S., Tofah, P., & Adenuga, O. M. (2021). Advances in non-destructive testing methods for weld integrity evaluation in high-capacity power infrastructure. *International Journal of Engineering and Modern Technology*, 7(1), 20-47. <https://doi.org/10.56201/ijemt.vol.7.no1.2021.pg20.47>
- Badmus, O., Jooda, D., & Anunagba, C. O. (2024). A privacy-by-design framework for role-based security architecture in Salesforce environments handling sensitive personal data. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3244-3254. <https://doi.org/10.62225/2583049X.2024.4.6.6243>
- Bola-Sadipe, D., Eboh, E. E., & Odihi, F. (2022). Advances in data-driven tax compliance and financial reporting systems in banking institutions. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(1), 366-403. <https://doi.org/10.32628/SHISRRJ247137>
- Borah, S., Aliliele, K. C., Rakshit, S., & Vajjhala, N. R. (2022). Applications of artificial intelligence in software testing. In P. K. Mallick et al. (Eds.), *Cognitive informatics and soft computing (Lecture Notes in Networks and Systems, Vol. 375, pp. 727-736)*. Springer. https://doi.org/10.1007/978-981-16-8763-1_60
- Borg, M., Runeson, P., & Ardö, A. (2014). Recovering from a decade: A systematic mapping of information retrieval approaches to software traceability. *Empirical Software Engineering*, 19, 1565-1616. <https://doi.org/10.1007/s10664-013-9255-y>
- Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., & Zafar, B. (2016). Financial education and the debt behavior of the young. *Review of Financial Studies*, 29(9), 2490-2522. <https://doi.org/10.1093/rfs/hhw006>

- Catal, C., & Mishra, D. (2013). Test case prioritization: A systematic mapping study. *Software Quality Journal*, 21, 445-478. <https://doi.org/10.1007/s11219-012-9181-z>
- Cleland-Huang, J., Gotel, O., & Zisman, A. (Eds.). (2012). *Software and Systems Traceability*. Springer. <https://doi.org/10.1007/978-1-4471-2239-5>
- Dada, T., Isiekwu, C. P., & Oluwo, K. (2021a). Advances in multinational cash flow consolidation and FX risk control using tiered treasury architecture. *Iconic Research and Engineering Journals*, 4(11), 601-620. <https://doi.org/10.64388/IREV4111-1714351>
- Daka, E., & Fraser, G. (2014). A survey on unit testing practices and problems. *IEEE International Symposium on Software Reliability Engineering Workshops*, 201-211. <https://doi.org/10.1109/ISSREW.2014.11>
- De Lucia, A., Fasano, F., Oliveto, R., & Tortora, G. (2007). Recovering traceability links in software artifact management systems using information retrieval methods. *ACM Transactions on Software Engineering and Methodology*, 16(4), Article 13. <https://doi.org/10.1145/1276933.1276934>
- Dogbatsey, E. A., Ebhojie, O., & Oyeleye, A. O. (2019). Reconciliation control and financial workflow redesign in emerging market organizations: A conceptual framework. *Zenodo*. <https://doi.org/10.5281/zenodo.20447103> [Originally numbered 3(5); journal name not captured in source.]
- Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2), 1-31. <https://doi.org/10.1257/app.6.2.1>
- Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and Software Technology*, 50(9-10), 833-859. <https://doi.org/10.1016/j.infsof.2008.01.006>
- Easterbrook, S., Singer, J., Storey, M., & Damian, D. (2008). Selecting empirical methods for software engineering research. In *Guide to Advanced Empirical Software Engineering* (pp. 285-311). Springer.
- Ekelemu, O. (2025a). Control room alarm management and human factors in large liquefied natural gas facilities: A comprehensive review of rationalization frameworks, operator response, and the elimination of standing and inhibited alarms. *International Journal of Engineering and Modern Technology*, 11(8), 352-396. <https://doi.org/10.56201/ijemt.vol.11.no8.2025.pg352.396>
- Ekelemu, O., & Akano, O. A. (2025b). Loss detection and integrity surveying in liquefaction header systems: Flare hunting, attribution and programme design for unintended flow to flare. *International Journal of Engineering and Modern Technology*, 11(8), 458-487. <https://doi.org/10.56201/ijemt.vol.11.no8.2025.pg458.487>
- Ekelemu, O. (2025b). Cross-shift information loss and the standardization of logbook handovers: A critical review of communication failure and continuity protocols in continuous processing operations. *International Journal of Engineering and Modern Technology*, 11(8), 397-424. <https://doi.org/10.56201/ijemt.vol.11.no8.2025.pg397.424>
- Eyetsemitan, R. A., Ambali, K. B., Oyeleye, A. O., & Fadayomi, O. (2023a). Change management in small business digital transformation: A systematic review and lean change adoption framework. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(6), 521-550. <https://doi.org/10.32628/GISRRJ236648>
- Eyetsemitan, R. A., Ambali, K. B., Oyeleye, A. O., & Fadayomi, O. (2023b). User acceptance testing in small business technology deployment: A structured validation framework for lean operational environments. *International Journal of Multidisciplinary Research and Growth Evaluation*, 4(6), 1512-1531. <https://doi.org/10.54660/IJMRGE.2023.4.6.1512-1531>

- Eze, F. I., Akinleye, O. K., & Anene, U. N. (2024b). Development of a risk-based accelerated approval decision model balancing speed, safety, and evidence quality in pharmaceutical regulatory science. *Research Journal of Pure Science and Technology*, 7(6), 120-161. <https://doi.org/10.56201/rjpst.v7.no6.2024.pg120.161>
- Eze, F. I., Anene, U. N., & Akinleye, O. K. (2022b). Design of an integrated regulatory readiness maturity model for pharmaceutical companies seeking global market entry. *International Journal of Health and Pharmaceutical Research*, 7(1), 77-112. <https://doi.org/10.56201/ijhpr.v7.no1.2022.pg77.112>
- Ezeako, R., Ibeh, A. V., Alugwe, S., & Awulu, I. (2024). Performance dashboards and data-driven decision-making in operations management: A conceptual framework for leadership visibility. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3520-3538. <https://doi.org/10.62225/2583049X.2024.4.6.6930>
- Fadayomi, O., Abolaji, T. O., Edivri, J., Ogbole, J. I., Okoruwa, P. O., & Akeju, B. (2019). Risk-based cybersecurity assurance and data availability: Limitations, advances and future research opportunities. *Iconic Research and Engineering Journals*, 2(12), 602-617. <https://doi.org/10.64388/IREV2I12-1713779>
- Fadayomi, O., Abolaji, T. O., Edivri, J., Ogbole, J. I., Okoruwa, P. O., & Akeju, B. (2019). Risk-based cybersecurity assurance and data availability limitations, advances and future research opportunities. *Iconic Research and Engineering Journals*, 2(12), 602–617. <https://doi.org/10.64388/IREV2I12-1713779>
- Fadayomi, O., Abolaji, T. O., Edivri, J., Ogbole, J. I., Okoruwa, P. O., & Akeju, B. (2019). Risk-based cybersecurity assurance and data availability: Limitations, advances and future research opportunities. *IRE Journals*, 2(12). <https://doi.org/10.64388/IREV2I12-1713779>
- Falemi, A., Akhigbe, R., & Akin-Oluyomi, O. T. (2024). A conceptual enterprise risk management model that integrates HR strategy with operational analytics. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(6), 1730-1753. <https://doi.org/10.54660/ijmrge.2024.5.6.1730-1753>
- Filani, O. M., Nnabueze, S. B., Ike, P. N., & Wedraogo, L. (2022). Real-time risk assessment dashboards using machine learning in hospital supply chain management systems. *International Journal of Multidisciplinary Evolutionary Research*, 3(1), 65-76. <https://doi.org/10.54660/IJMER.2022.3.1.65-76>
- Filani, O. M., Nnabueze, S. B., Sakyi, J. K., & Okojie, J. S. (2023). Scenario-based financial modelling for enhancing strategic decision-making and organizational long-term planning. *Journal of Frontiers in Multidisciplinary Research*, 4(2), 251–265. <https://doi.org/10.54660/JFMR.2023.4.2.251-265>
- Filani, O. M., Nnabueze, S. B., Sakyi, J. K., & Okojie, J. S. (2023). Scenario-based financial modelling for enhancing strategic decision-making and organizational long-term planning. *Journal of Frontiers in Multidisciplinary Research*, 4(2), 251-265. <https://doi.org/10.54660/JFMR.2023.4.2.251-265>
- Fobellah, A. N. (2025a). Cognitive biases in financial decision-making: Implications for audit and risk management in large corporations: A conceptual review. *International Journal of Science and Research Archive*, 16(2), 17-22. <https://doi.org/10.30574/ijrsra.2025.16.2.2273>
- Fobellah, A. N. (2025c). Unintended consequences of financial regulations: A study of corporate compliance burdens: A conceptual review. *International Journal of Science and Research Archive*, 16(2), 29-34. <https://doi.org/10.30574/ijrsra.2025.16.2.2275>
- Gotel, O. C. Z., & Finkelstein, A. C. W. (1994). An analysis of the requirements traceability problem. *IEEE International Conference on Requirements Engineering*, 94-101. <https://doi.org/10.1109/ICRE.1994.292398>
- 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>

- Ilodigwe, L., & Adesemoye, A. C. (2024a). Advances in real-world evidence, predictive analytics, and data-driven decision making for improving health system performance and patient outcomes. *International Journal of Health and Pharmaceutical Research*, 9(5), 174-197. <https://doi.org/10.56201/ijhpr.v9.no5.2024.pg174.197>
- Ilodigwe, L., & Adesemoye, A. C. (2025a). Advances, risks, and implementation challenges of artificial intelligence as a force multiplier for clinical decision making and health system efficiency. *International Journal of Medical Evaluation and Physical Report*, 9(7), 165-185. <https://doi.org/10.56201/ijmepr.v9.no7.2025.pg165.185>
- Ilodigwe, L., & Adesemoye, A. C. (2024b). Beyond technology: A strategic framework for building sustainable and resilient health systems through organizational strategy, operational excellence, and workforce capability. *International Journal of Medical Evaluation and Physical Report*, 8(6), 299-319. <https://doi.org/10.56201/ijmepr.v8.no6.2024.pg299.319>
- ISO/IEC/IEEE. (2018). ISO/IEC/IEEE 29119-2 Software and systems engineering: Software testing processes. International Organization for Standardization.
- ISO/IEC/IEEE. (2018). ISO/IEC/IEEE 29148 Systems and software engineering: Life cycle processes and requirements engineering. International Organization for Standardization.
- Jimoh, H. O., & Ahmed, M. O. (2024). Analyzing Network Time Protocol (NTP) based amplification DDoS attack and its mitigation techniques. *Journal of Digital Innovations and Contemporary Research in Science, Engineering and Technology*, 12(2), 17-24. <https://doi.org/10.22624/AIMS/DIGITAL/V11N2P2x>
- Jones, B. C., & Ominyi, M. (2021). Developing a conceptual framework linking West African agricultural value chains to household economic stability. *International Journal of Economics and Financial Management*, 6(2), 89-114. <https://doi.org/10.56201/ijefm.v6.no2.2021.pg89.114>
- Kaner, C., Falk, J., & Nguyen, H. Q. (1999). *Testing Computer Software* (2nd ed.). Wiley.
- Kirchler, E. (2007). *The Economic Psychology of Tax Behaviour*. Cambridge University Press.
- Komi, N. M. (2023). When accurate forecasts are not fair forecasts: Deep learning load prediction and equity across districts. Shodhshauryam, *International Scientific Refereed Research Journal*, 6(6), 476-546. <https://doi.org/10.32628/SHISRRJ247102>
- Kumuyi, O., Uzoka, E., Akeju, B., & Ozowara, D. E. (2024b). Framework for privacy-focused digital identity verification supporting financial inclusion in Africa. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3150-3162. <https://doi.org/10.62225/2583049X.2024.4.6.6005>
- Ladapo, O. O., Jooda, D., Dosunmu, A. A., & Abolaji, T. O. (2023b). Systematic literature review on security access control policies and techniques based on privacy requirements in a BYOD environment. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2890-2904. <https://doi.org/10.62225/2583049X.2023.3.6.6206>
- Li, Z., Harman, M., & Hierons, R. M. (2007). Search algorithms for regression test case prioritization. *IEEE Transactions on Software Engineering*, 33(4), 225-237. <https://doi.org/10.1109/TSE.2007.38>
- Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics and Finance*, 14(4), 332-368. <https://doi.org/10.1017/S1474747215000232>
- Mairiza, D., Zowghi, D., & Nurmuliani, N. (2010). An investigation into the notion of non-functional requirements. *ACM Symposium on Applied Computing*, 311-317. <https://doi.org/10.1145/1774088.1774153>
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Oluoha, O. M. (2018). A conceptual framework for legal and ethical risk modeling in enterprise data protection governance systems. *Iconic Research and Engineering Journals*, 2(2), 207-226. <https://doi.org/10.64388/IREV2I2-1714911>

- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2021a). A conceptual framework for risk based business intelligence architecture in financial technology platforms. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 731-746. <https://doi.org/10.54660/IJMRGE.2021.2.6.731-746>
- Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2022a). A review of data protection impact assessment models in multi cloud financial infrastructure systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1), 589-623. <https://doi.org/10.32628/CSEIT25442>
- Michael, O. N., & Ogunsola, O. E. (2025c). Assessing the role of artificial intelligence in transforming decision-making across modern agricultural systems. *Engineering and Technology Journal*, 10(12), 8027-8048. <https://doi.org/10.47191/etj/v10i12.06>
- Michael, O. N., & Ogunsola, O. E. (2021d). Investigating the contribution of climate-smart agricultural practices to household food security and income stability. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(2), 209-219. <https://doi.org/10.54660/IJMRGE.2020.1.2.209-219>
- Miller, M., Reichelstein, J., Salas, C., & Zia, B. (2015). Can you help someone become financially capable? *World Bank Research Observer*, 30(2), 220-246. <https://doi.org/10.1093/wbro/lkv009>
- Myers, G. J., Sandler, C., & Badgett, T. (2011). *The Art of Software Testing* (3rd ed.). Wiley.
- Netemeyer, R. G., Warmath, D., Fernandes, D., & Lynch, J. G. (2018). How am I doing? Perceived financial well-being. *Journal of Consumer Research*, 45(1), 68-89. <https://doi.org/10.1093/jcr/ucx109>
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2024b). Securing agentic AI enterprise workflows against prompt injection, tool poisoning, memory manipulation, and excessive agency threats. *World Journal of Innovation and Modern Technology*, 8(6), 217-248. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg217.248>
- Obogo, S. F., Nwafor, M. I., & Ozobu, C. O. (2023). Conceptual leadership model for safety culture development in construction and engineering projects. *International Journal of Scientific Research in Civil Engineering*, 7(6), 121-153. <https://doi.org/10.32628/IJSRCE237554>
- Obogo, S. F., Ozobu, C. O., & Uduokhai, D. O. (2019a). Advances in leadership driven safety culture transformation in large construction workforces. *Iconic Research and Engineering Journals*, 3(5), 507-523. <https://doi.org/10.64388/IREV3I5-1715497>
- Obogo, S. F., Ozobu, C. O., & Uduokhai, D. O. (2019b). Development of a construction safety risk governance model for multi contractor high rise projects. *Iconic Research and Engineering Journals*, 2(9), 502-522. <https://doi.org/10.64388/IREV2I9-1715495>
- Obogo, S. F., Uduokhai, D. O., & Ozobu, C. O. (2019). Systematic review of hazard identification and risk control practices in urban construction projects. *Iconic Research and Engineering Journals*, 2(12), 666-681. <https://doi.org/10.64388/IREV2I12-1715496>
- Obriki, O. D., & Arumosoye, O. M. (2019). A conceptual framework linking management safety walkthrough frequency and coverage to safety culture outcomes in mega projects. *Iconic Research and Engineering Journals*, 2(8), 355-374. <https://doi.org/10.64388/IREV2I8-1714416>
- Obriki, O. D., & Arumosoye, O. M. (2020). Conceptual framework for human error causation in high-risk construction and industrial activities. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 715-727. <https://doi.org/10.54660/IJMRGE.2020.1.5.715-727>

- Obriki, O. D., Arumosoye, O. M., & Ozobu, C. O. (2025). Conceptual model linking leading safety signals to sustained injury-free project performance. *International Journal of Scientific Research in Humanities and Social Sciences*, 2(3), 233-254. <https://doi.org/10.32628/IJSRHSS252342>
- Ogbete, J. C., Aminu-Ibrahim, A. Y., & Ambali, K. B. (2022). Design standards and operational planning frameworks for scalable blood collection networks. *Gyanshauryam, International Scientific Refereed Research Journal*, 5(6), 267-300. <https://doi.org/10.32628/GISRRJ225635>
- Ogbete, J. C., Aminu-Ibrahim, A. Y., & Ambali, K. B. (2023). Lifecycle performance evaluation of purpose built diagnostic laboratories supporting long term healthcare delivery. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2605-2621. <https://doi.org/10.62225/2583049X.2023.3.6.5698>
- Ogbete, J. C., Aminu-Ibrahim, A. Y., & Ambali, K. B. (2018). Optimizing laboratory spatial planning strategies to improve diagnostic accuracy, safety, and clinical throughput. *Iconic Research and Engineering Journals*, 2(1), 87-113. <https://doi.org/10.64388/IREV117-1713587>
- Ogbole, J. I., Okoruwa, P. O., Babatope, O. M., & Oyewole, T. (2021). Developing an integrated data visualization model for continuous business performance monitoring and optimization. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 441-467. <https://doi.org/10.32628/IJSRCSEIT>
- Ogbole, J. I., Okoruwa, P. O., Babatope, O. M., & Oyewole, T. (2021). Developing an integrated data visualization model for continuous business performance monitoring and optimization. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 441-467. <https://doi.org/10.32628/IJSRCSEIT>
- Oluwadele, J. F., Samuel, O., Tawose, O. M., Ekeocha, A. H., & Adika, A. O. (2025). Evaluation of alternative energy sources to replace maize in Marshall broiler diets: Effects on growth performance, meat quality, and serum biochemistry. *EUREKA: Life Sciences*. <https://doi.org/10.21303/2504-5695.2025> \[Author list truncated in source; verify full authorship.\]
- Oluwo, K., Bola-Sadipe, D., & Aliliele, C. (2025). Conceptual framework for data analytics driven financial decision making in banking sector organizations. *Shodhshauryam, International Scientific Refereed Research Journal*, 8(5), 259-302. <https://doi.org/10.32628/SHISRRJ258529>
- Ominyi, M., Anichukwueze, C. C., & Uzougbo, N. S. (2025a). A conceptual framework for multi-agent AI quality control in the review of regulated documents. *International Journal of Social Sciences and Management Research*, 11(8), 534-557. <https://doi.org/10.56201/ijssmr.vol.11no8.2025.pg534.557>
- Ominyi, M., Anichukwueze, C. C., & Uzougbo, N. S. (2024b). A systematic review of organizational readiness for the EU Artificial Intelligence Act in multinational enterprises. *World Journal of Innovation and Modern Technology*, 8(6), 185-203. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg185.203>
- Onche, V. O., Adegbite, M. P., & Ogbonna, C. S. (2023). BYOD adoption and information-security risk in university administrative offices: Evidence from Nigerian federal universities. *International Journal of Multidisciplinary Research and Growth Evaluation*, 4(6), 1549-1563. <https://doi.org/10.54660/IJMRGE.2023.4.6.1549-1563>
- Onwuzulike, O. C., Uzundu, N. C., & Joseph, O. B. (2024). Impact of strategic management education on professional development: A longitudinal study. *World Journal of Advanced Research and Reviews*, 23(3), 323-337. <https://doi.org/10.30574/wjarr.2024.23.3.2672>
- Oyeleye, A. O., Dogbatsey, E. A., & Ebhojie, O. (2025). Embedding automated close controls in public sector ERP systems: A conceptual model for audit readiness and financial reporting quality. *Zenodo*. <https://doi.org/10.5281/zenodo.20097950> \[Originally numbered 2(1), 110-130; journal name not captured in source.\]

- Page, M. J., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Ramesh, B., & Jarke, M. (2001). Toward reference models for requirements traceability. *IEEE Transactions on Software Engineering*, 27(1), 58-93. <https://doi.org/10.1109/32.895989>
- Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14, 131-164. <https://doi.org/10.1007/s10664-008-9102-8>
- Sakyi, J. K., Filani, O. M., Nnabueze, S. B., Okojie, J. S., & Ogedengbe, A. O. (2022b). Developing KPI frameworks to enhance accountability and performance across large-scale commercial organizations. *Journal of Frontiers in Multidisciplinary Research*, 3(2), 81-93. <https://doi.org/10.54660/IJFMR.2022.3.2.81-93>
- Sanni, J. O., Atima, M. E., & Attah, A. (2022). Systematic review of attribution modeling methods resolving bias in multi-touch journeys. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(1), 304-334. <https://doi.org/10.32628/SHISRRJ247135>
- Sanni, J. O., & Atima, M. E. (2021b). Business intelligence dashboard frameworks resolving executive visibility gaps in strategic marketing governance. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 633-646. <https://doi.org/10.54660/IJMRGE.2021.2.6.633-646>
- Sanni, J. O., Iwuanyanwu, U. A., Essien, M. A., & Attah, A. (2023). Lifecycle-aware marketing automation using federated learning for secure cross-organizational data management. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(6), 337-364. <https://doi.org/10.32628/GISRRJ236642>
- Sanni, J. O., Iwuanyanwu, U. A., & Essien, M. A. (2025). Problem-oriented process mining for auditable marketing automation lifecycle control. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 1933-1947. <https://doi.org/10.62225/2583049X.2025.5.6.5650>
- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2023). A conceptual framework for retail asset valuation: Integrating quantitative pricing and multi-unit financial performance prediction. *Journal of Accounting and Financial Management*, 9(12), 218-250. <https://doi.org/10.56201/jafm.v9.no12.2023.pg218.250>
- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2024b). Real-time KPI tracking systems: A review of automated performance monitoring and data-driven decision making. *World Journal of Innovation and Modern Technology*, 8(6), 247-281. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg247.281>
- Uchechi, N., Ominyi, M., Anichukwueze, C. C., & Uzougbo, N. S. (2022). Advances in quantitative evaluation models for measuring corporate diversity and inclusion investments. *World Journal of Finance and Investment Research*, 6(2), 56-101. <https://doi.org/10.56201/wjfir.v6.no2.2022.pg56.101>
- Uchechi, N., Ominyi, M., & Jones, B. C. (2018). Procedural standardization and workflow optimization in high-volume dispute resolution clinics: A systematic review. *Iconic Research and Engineering Journals*, 2(6). <https://doi.org/10.64388/IREV2I6-1722789>
- Wedraogo, L., & Sanni, J. O. (2024). Machine learning models addressing uncertainty in cross-channel campaign performance forecasting accuracy. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 2875-2890. <https://doi.org/10.62225/2583049X.2024.4.6.5648>
- Wieringa, R. J. (2014). *Design Science Methodology for Information Systems and Software Engineering*. Springer. <https://doi.org/10.1007/978-3-662-43839-8>

World Bank. (2014). Global Financial Development Report 2014: Financial Inclusion. World Bank. <https://doi.org/10.1596/978-0-8213-9985-9>

Xiao, J. J., & O'Neill, B. (2016). Consumer financial education and financial capability. *International Journal of Consumer Studies*, 40(6), 712-721. <https://doi.org/10.1111/ijcs.12285>

Yusuf, H. D., & Abdulrazaq, K. (2018). A conceptual framework for strengthening diagnostic quality systems in resource-limited clinical laboratories. *Iconic Research and Engineering Journals*, 1(11), 145-179. <https://doi.org/10.64388/IREV1I11-1723090>

Zhu, H., Hall, P. A. V., & May, J. H. R. (1997). Software unit test coverage and adequacy. *ACM Computing Surveys*, 29(4), 366-427. <https://doi.org/10.1145/267580.267590>