Volume 4, Issue 4, April - 2026 **ISSN(E):** 2938-3773

DEVELOPMENT AND VALIDATION OF DIAGNOSTIC INSTRUMENTS FOR ASSESSING CYBERSECURITY LITERACY AMONG VOCATIONAL EDUCATION STUDENTS

Abdunazarov Baxodir Abduazizovich
Independent Researcher at the
Institute for Development of Vocational Education

Abstract
The contemporary landscape of vocational education and training is characterized by accelerating digitalization, where students entering technical, agricultural, and service-sector professions must navigate increasingly complex cyber-physical environments in which a single uninformed click can compromise organizational data, financial assets, or critical infrastructure. This article addresses the critical absence of psychometrically validated, context-sensitive diagnostic instruments for measuring cybersecurity literacy among vocational education students in Uzbekistan, and proposes a comprehensive multi-component diagnostic framework as a methodological mechanism for evidence-based competency assessment.

Keywords: Vocational education management, cybersecurity literacy, competency-based assessment, triarchic intelligence model, psychometric validation, multi-method diagnostics, digital readiness, Uzbekistan.

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3

Introduction

In the context of accelerating Industry 4.0 transformations and the global digitalization of labor markets, the management of vocational education institutions (VEIs) faces a dual challenge: equipping students with technical competencies aligned with rapidly evolving occupational standards while simultaneously cultivating transversal digital and cybersecurity competencies that have become non-negotiable in nearly every modern profession. Traditional vocational assessment in Uzbekistan operates on a vertical hierarchy where curricular content, learning outcomes, and evaluation instruments flow from ministerial standards down to college-level testing, often with minimal feedback from industry stakeholders or empirical psychometric calibration. While this model ensures procedural uniformity across the national vocational education system, it suffers from significant "diagnostic blindness" — an inability to capture the complex, multidimensional nature of cybersecurity literacy as it manifests in real workplace contexts.

From a methodological perspective, this diagnostic blindness manifests as the predominance of multiple-choice knowledge tests that measure recognition of technical terminology while leaving behavioral readiness, threat perception, and ethical reasoning entirely uncharted. From an organizational perspective, the traditional model produces fragmented assessment data that cannot be aggregated into actionable institutional indicators or used to benchmark student readiness against international frameworks such as DigComp 2.2 or the NIST NICE Workforce

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3

Framework. The result is a paradox: high formal pass rates in informatics modules coexist with systematically documented vulnerabilities of vocational graduates to phishing, social engineering, and unsafe data-handling behaviors during their first employment positions. This systemic dysfunction necessitates a transition to a "Comprehensive Diagnostic Framework" — a psychometrically grounded, multi-method approach that captures cybersecurity literacy as an integrated cognitive, behavioral, and dispositional construct rather than a single recall-based outcome.

The theoretical basis for applying multi-dimensional diagnostic measurement to vocational cybersecurity education is derived from Glaser's theory of knowledge-rich cognition, which positions structured domain expertise as the substrate of all higher-order reasoning and adaptive performance. Glaser defined competence as a complex interaction of structured knowledge and adaptive procedural skill exercised under contextual constraints. In the context of cybersecurity literacy, this theoretical position translates to a measurement target that cannot be reduced to factual recall — a multidimensional construct comprising declarative knowledge of threats and defenses, procedural know-how for incident response, and adaptive judgment under conditions of social engineering and information uncertainty. Sternberg's triarchic theory further reinforces this position by distinguishing analytic, practical, and creative components of intelligent performance, each of which must be independently sampled by a defensible diagnostic instrument.

Unlike simple knowledge tests (e.g., terminology multiple-choice items administered in isolation), a comprehensive diagnostic instrument represents a higher order of measurement characterized by the following theoretical attributes:

- **Construct Multidimensionality:** The instrument simultaneously operationalizes at least three distinct latent traits — cybersecurity knowledge, behavioral self-efficacy in threat scenarios, and security-oriented attitudes. Each subscale must demonstrate independent reliability (Cronbach's alpha not less than 0.75) while contributing to an interpretable total score, in line with Kholodnaya's psychometric tradition that treats intelligence and competence as structured rather than monolithic.

- **Ecological Validity:** Items are anchored in scenarios drawn from authentic vocational occupational contexts — welding workshops with networked CNC machines, accounting offices using cloud-based ERP systems, agricultural technicians operating IoT sensor networks. This anchoring ensures that performance on the instrument predicts performance in genuine workplace situations rather than only in classroom test conditions.

- **Diagnostic Granularity:** Rather than producing a single pass-or-fail outcome, the instrument generates a profile across sub-competencies, enabling pedagogical interventions targeted at specific developmental gaps (e.g., low scores in social engineering recognition trigger scenario-based remedial modules) rather than blanket remediation that wastes instructional time.

Thus, the adoption of multi-method diagnostic instrumentation is not merely a technical refinement of testing procedures but a fundamental shift in the epistemology of vocational

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3

assessment. It moves the system from a "summative gatekeeping" model to a "formative developmental" model, where measurement outcomes feed directly into individualized learning trajectories rather than serving as terminal credentialing decisions. In the developmental model, the diagnostic instrument is not the final filter at the end of the pipeline but a continuous instrument of curricular adjustment, professional development, and strategic policy planning at the institutional and ministerial levels.

The transition to comprehensive diagnostic measurement requires a radical restructuring of how cybersecurity competencies are operationalized at the item, instrument, and reporting levels. The current single-instrument approach is methodologically insufficient because no individual measurement modality — whether knowledge tests, self-report scales, or performance tasks — can simultaneously achieve high content validity, ecological validity, and scalable administration. Therefore, the core of the proposed methodological mechanism is the establishment of a Tripartite Diagnostic Battery coordinated by an institutional Assessment Quality Council with executive rather than merely advisory authority over item development, item retirement, and validation studies.

The Council must not function as a peripheral committee but as the operational owner of the entire assessment pipeline, with binding authority over content blueprints and item-bank governance. Structurally, the Tripartite Diagnostic Battery must represent the three foundational domains of cybersecurity literacy:

1. **Cognitive Domain:** A 30-item criterion-referenced knowledge test covering threat taxonomies, defensive measures, legal and ethical norms, and incident response procedures (calibrated through Item Response Theory to ensure balanced difficulty distribution and adequate item discrimination across the ability range).

2. **Behavioral Domain:** A 12-scenario simulation module presenting realistic phishing emails, suspicious USB devices, password-reset requests, and data-disclosure dilemmas (scored through expert-rated decision trees to capture procedural readiness in time-constrained, ambiguous situations characteristic of actual workplace incidents).

3. **Affective Domain:** A 20-item Likert-type attitudinal scale measuring security self-efficacy, perceived threat severity, and intention to comply with cybersecurity protocols (validated through confirmatory factor analysis to capture the dispositional substrate that mediates between knowledge and behavior).

The methodological innovation lies in the standardized aggregation procedure that combines these heterogeneous data streams into a single Cybersecurity Literacy Profile rather than collapsing them into an interpretively impoverished total score. This ensures that the final diagnostic output is "backward engineered" from the multidimensional construct definition rather than artificially compressed into a unitary number that obscures the very profile differences pedagogues need to act upon. A complementary peer-review subcommittee within the Assessment Quality Council oversees periodic item revision based on classroom feedback, validity evidence, and emerging threat landscapes that quickly render previously valid items obsolete.

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3

To address scalability constraints across a geographically dispersed vocational education system, the framework introduces shared computer-based testing centers at the regional level, eliminating the redundancy of each individual institution maintaining isolated testing infrastructure of variable quality. These centers, equipped with secure browser environments and centralized item banks, increase utilization efficiency, reduce per-student administration costs, and ensure standardization of test conditions across institutions of differing technological maturity.

The economic and methodological viability of the diagnostic framework hinges on a robust validation pipeline that justifies the resource investment through demonstrated measurement quality and policy utility. The proposed economic mechanism operates on the principle of Multi-Stage Psychometric Validation, which integrates three distinct validation streams:

- **Content Validation Stream (Foundational Level):** Covers the systematic mapping of items to a competency blueprint derived from national vocational education standards, the DigComp 2.2 framework, and structured industry interviews. This is operationalized through expert panels composed of eight to twelve cybersecurity practitioners and pedagogues who rate item relevance using Lawshe's Content Validity Ratio, with a retention threshold of 0.62.

- **Construct Validation Stream (Targeted Level):** This is conducted through pilot administration to a stratified sample of three hundred to five hundred vocational education students drawn proportionally from technical, agricultural, and service-sector colleges. The cluster entity responsible for analysis applies confirmatory factor analysis and inter-scale correlation matrices to test whether the empirical data align with the three-factor theoretical structure underlying the Tripartite Diagnostic Battery.

- **Predictive Validation Stream (Commercial Level):** The diagnostic battery is administered to graduating students, and their scores are subsequently correlated with workplace cybersecurity incident records and employer performance ratings collected six months post-graduation. This stream provides the evidentiary link between assessment outcomes and professional consequences, justifying the investment for both the labor market and the funding ministry.

However, sustained psychometric quality requires a robust system of Methodological Incentives that motivates ongoing item refinement and discourages the institutional drift that allows test instruments to decay through unchecked exposure and changing threat landscapes. These incentives include institutional research grants for instrument developers, formal recognition of validation studies as equivalent to peer-reviewed publications in academic promotion criteria, and the introduction of an innovative methodological instrument: the concept of “Living Item Banks” — dynamically updated repositories where exposed or psychometrically degraded items are systematically retired and replaced with new ones reflecting emerging threats, ensuring that the instrument never becomes obsolete or compromised.

This Living Item Bank concept reduces the long-term measurement risk for vocational education institutions, which would otherwise face periodic instrument decay and the costly

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3

necessity of full-scale re-validation, and guarantees a stable methodological return on investment for the Agency for Vocational Education, which can rely on consistent national-level data for policy decisions across years and successive student cohorts. The economic logic mirrors the actuarial logic of insurance pools: distributed item-bank costs across the system stabilize the cost-per-administration while raising aggregate measurement quality.

A crucial component of the diagnostic mechanism is the digitalization of administration, scoring, and reporting through a unified Cybersecurity Diagnostic Information System. The system aggregates anonymized response-level data from all participating vocational education institutions, automatically computes subscale and composite scores, performs differential item functioning analyses across demographic and regional groups, and generates institutional dashboards that visualize cohort-level competency profiles. Decision-making at the curricular level is informed by real-time aggregate diagnostics: when more than thirty percent of a cohort demonstrates a behavioral subscale gap, the system triggers a notification to the corresponding curriculum committee with recommended remedial modules. This data-driven approach minimizes resource waste associated with "blanket remediation" and "diagnostic latency" that characterize the current paper-based assessment cycle, in which actionable feedback typically arrives months after the relevant instructional window has closed.

The implementation of the Tripartite Diagnostic Battery and its supporting psychometric mechanisms is projected to yield transformative results for the quality of cybersecurity preparation in the vocational education system. Based on comparative analyses of international diagnostic frameworks (e.g., Finland's competence-based assessment model in vocational schools, Germany's dual-system performance examinations), the following outcomes are anticipated:

- **Measurement Reliability:** Cronbach's alpha coefficients in the range of 0.85 to 0.90 across the three subscales, with test-retest reliability of approximately 0.80 over a four-week interval, ensuring that diagnostic outputs are stable enough to inform individual-level pedagogical decisions and longitudinal cohort tracking.

- **Predictive Validity:** Correlations of 0.55 to 0.65 between battery scores and workplace cybersecurity incidents during the first six months of employment, demonstrating that the instrument identifies "deadstock graduates" — those formally credentialed but behaviorally unprepared — with sufficient sensitivity to enable targeted pre-graduation interventions rather than reactive post-employment remediation.

- **Pedagogical Impact:** A projected reduction of twenty to thirty percent in workplace cybersecurity incidents reported by employers of vocational education graduates within three years of full implementation, alongside a measurable shift in instructional time toward scenario-based learning that the diagnostic profile reveals to be deficient relative to declarative-knowledge instruction.

While the theoretical argument for a comprehensive diagnostic framework is compelling, the practical implementation faces significant "methodological friction" arising from the gap between psychometric ideals and the operational realities of a heterogeneous vocational

education system. The discussion highlights three primary risks that must be managed through deliberate institutional design:

- **Item Compromise Risk:** In a system administering identical items across hundreds of institutions, item exposure and informal sharing among student communities can rapidly degrade item parameters and compromise score interpretability. The consequence is inflated mean scores that mask genuine competency deficits and undermine the predictive validity of the instrument. Mitigation requires rotating item banks, item-level usage monitoring with automatic exposure thresholds, and the rapid retirement of overexposed items through the Living Item Bank mechanism.

- **Pedagogical Distortion Risk:** When high-stakes diagnostic outputs influence institutional rankings or funding allocations, instructors may engage in "teaching to the test" — narrowly preparing students for specific item formats rather than building generative cybersecurity competence. The consequence is artifactual score gains without corresponding behavioral readiness, producing the very deadstock graduates the framework is designed to detect. Mitigation requires the systematic use of unannounced parallel forms, scenario-based items resistant to memorization, and a deliberate decoupling of diagnostic outputs from punitive institutional consequences.

- **Equity Risk:** Differential access to digital infrastructure and prior computing experience can produce systematic score disparities between urban and rural vocational education institutions that reflect environmental rather than competency differences. The consequence is the reinforcement of existing educational inequities through diagnostic labeling, contradicting the developmental purpose of the framework. Mitigation requires routine differential item functioning analyses, stratified score interpretation guidelines, and deliberate over-sampling of rural and remote cohorts during validation studies to ensure that item parameters generalize across the full system.

In conclusion, the development of a Tripartite Diagnostic Battery for cybersecurity literacy is not merely a technical assessment exercise but a strategic necessity for aligning vocational education in Uzbekistan with the demands of a digitally transformed economy and an increasingly hostile cyber threat landscape. It represents a shift from "credential-counting" to "competency-evidence" paradigms, where institutional and individual decisions are anchored in psychometrically defensible data rather than in presumed curricular coverage. The construct-based organizational mechanism ensures that measurement targets remain faithful to the multidimensional nature of cybersecurity literacy, while the multi-stage validation mechanism ensures that assessment investments yield reliable, predictive, and equitable outcomes. Together, these mechanisms transform diagnostic measurement from an administrative formality into a generative engine for continuous curricular and pedagogical improvement.

For the Institute for Development of Vocational Education and the Agency for Vocational Education, the following strategic actions are recommended to operationalize this model:

4. **Establish a National Assessment Quality Council:** Convene a permanent body comprising psychometricians, cybersecurity practitioners, vocational education pedagogues, and industry representatives (e.g., from IT firms, the banking sector, and operators of critical infrastructure) charged with overseeing item development, validation studies, and ongoing item bank maintenance.

5. **Pilot the Tripartite Diagnostic Battery in Selected Institutions:** Conduct a phased pilot in fifteen to twenty colleges across diverse specialization profiles and regional contexts, using the results to refine items, scoring rubrics, and aggregation algorithms before national rollout.

6. **Integrate Diagnostic Outputs into the Credit-Module Curriculum:** Use student-level diagnostic profiles to inform individualized learning pathways within the credit-module system (e.g., assigning targeted micro-modules in social engineering recognition for students with low behavioral subscale scores), thereby converting diagnostics into pedagogical action.

1. **Develop Capacity in Modern Psychometric Methods:** Establish a continuous professional development program for vocational education researchers in measurement methods including Item Response Theory, confirmatory factor analysis, and differential item functioning, ensuring sustainable national capacity for instrument maintenance and the principled adoption of future diagnostic technologies.

Ultimately, the success of cybersecurity diagnostic measurement in vocational education depends on a deeper principle — the recognition that what is measured shapes what is taught, and what is taught shapes what graduates do under pressure. Only through methodologically rigorous, ethically grounded, and pedagogically generative diagnostics can the vocational education system in Uzbekistan become a true driver of national digital resilience and graduate workforce readiness.

References:

1. O‘zbekiston Respublikasining “Ta’lim to‘g‘risida”gi O‘RQ-637-son Qonuni. 2020-yil 23-sentyabr // Qonunchilik ma’lumotlari milliy bazasi, 24.09.2020-y., 03/20/637/1313-son.
2. O‘zbekiston Respublikasi Prezidentining 2022-yil 28-yanvardagi PF-60-son “2022–2026-yillarga mo‘ljallangan Yangi O‘zbekistonning taraqqiyot strategiyasi to‘g‘risida”gi Farmoni.
3. Pozilova Sh.X. Raqamli jamiyat muhitida professional ta’lim o‘qituvchilarini kasbiy kreativligini rivojlantirishning didaktik asoslari va metodikasi: ped. fan. d-ri (DSc) diss. – Toshkent, 2024. – 261 b.
4. Kadirov X.Sh. Bo‘lajak kasb-ta’limi o‘qituvchilarida mediakompetentlikni rivojlantirish texnologiyasi: ped. fan. d-ri (DSc) diss. – Toshkent, 2020. – 237 b.
5. Karimova N.N. Paradigmal yondashuv asosida professional ta’lim pedagoglarining kasbiy pedagogik kompetentligini uzluksiz rivojlantirish: ped. fan. d-ri (DSc) diss. – Toshkent, 2023. – 254 b.

6. Raximov Z.T. Innovatsion yondashuv asosida bo‘lajak kasb ta’limi o‘qituvchilarining o‘quv-bilish kompetentligini rivojlantirish texnologiyasi: ped. fan. d-ri (DSc) diss. – Toshkent, 2021. – 233 b.
7. Glaser, R. Education and thinking: The role of knowledge // American Psychologist. – 1984. – Vol. 39 (2). – P. 93–104.
8. Sternberg, R.J., Grigorenko, E.L. Teaching for successful intelligence. – Arlington Heights, IL: Skylight Training and Publishing Inc., 2000.
9. Cedefop. Terminology of European education and training policy: A selection of 100 key terms. – Luxembourg: Publications Office of the European Union, 2008. – 248 p.
10. Зимняя, И.А. Ключевые компетенции — новая парадигма результата образования // Высшее образование сегодня. — 2003. — № 5. — С. 34–42.
11. Холодная, М.А. Психология интеллекта: парадоксы исследования. — СПб.: Изд-во Питер, 2002. — 272 с.

Web of Discoveries: Journal of Analysis and Inventions
webofjournals.com/index.php/3