

A DATA-DRIVEN AND THEORY-GUIDED FRAMEWORK FOR DEVELOPING AND VALIDATING HUMAN-ROBOT COLLABORATION TRAINING MODULES FOR THE CONSTRUCTION WORKFORCE

SUBMITTED: May 2026

PUBLISHED: August 2026

EDITOR: Robert Amor

DOI: [10.36680/j.itcon.2026.040](https://doi.org/10.36680/j.itcon.2026.040)

Ebenezer Olukanni, Ph.D. Candidate

Myers Lawson School of Construction, Virginia Tech, USA

<https://orcid.org/0000-0002-6086-5888>

ooebenezer@vt.edu

Abiola Akanmu, Professor

Myers Lawson School of Construction, Virginia Tech, USA

<https://orcid.org/0000-0001-9145-4865>

abiola@vt.edu

Houtan Jebelli, Assistant Professor

Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, USA

<https://orcid.org/0000-0003-4786-7616>

hjebelli@illinois.edu

SUMMARY: The growing integration of robotics and artificial intelligence in construction is reshaping project execution and creating new requirements for workforce capabilities in human-robot collaboration (HRC). However, construction organizations and academic institutions lack systematic and scalable approaches for translating HRC competencies into structured training systems aligned with workforce development and management needs. This study presents a data-driven and theory-guided framework for developing and validating HRC training modules to support construction workforce readiness. The framework integrates natural language processing-based competency augmentation, an ADDIE-driven instructional design process, and Delphi-based expert validation into a unified and reproducible workflow. An initial set of HRC competencies derived from prior literature was augmented using industry data, resulting in a validated framework of 50 HRC competencies across knowledge, skills, and abilities. These competencies were cognitively classified using Revised Bloom's Taxonomy and translated into measurable learning objectives through a structured, theory-constrained pipeline. Seven training modules were developed to reflect progressive competency development, from foundational robotics knowledge to system-level reasoning, safety, and performance evaluation. A two-round Delphi study with academic and industry experts established consensus on the relevance of competencies, module-competency alignment, and cognitive classification, while identifying challenges in operationalizing higher-order and socio-cognitive competencies into measurable outcomes. The findings demonstrate the robustness and practical relevance of the proposed framework for workforce training design. This study contributes to engineering management by providing a scalable methodology for translating HRC competencies into validated training modules and by demonstrating that assessment requirements differ systematically across knowledge, skill, and ability competencies, thereby supporting workforce planning, training standardization, assessment design, and construction robotics adoption.

KEYWORDS: Human-robot collaboration (HRC), Instructional design, Competency-based training, Construction robotics, Natural language processing.

REFERENCE: Olukanni, E., Akanmu, A., & Jebelli, H. (2026). A data-driven and theory-guided framework for developing and validating human-robot collaboration training modules for the construction workforce. *Journal of Information Technology in Construction (ITcon)*, 31, 939-969. <https://doi.org/10.36680/j.itcon.2026.040>

COPYRIGHT: © 2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The construction industry is undergoing significant technological transformations as robotics, automation, and artificial intelligence (AI) are increasingly integrated into field operations, off-site production, inspection, monitoring, layout, material handling, and other construction processes. Robotics and AI are being explored as responses to persistent challenges such as labor shortages, safety risks, productivity constraints, quality variability, and rising project complexity. Prior research has shown that construction automation is expanding beyond isolated mechanization toward more advanced robotic systems capable of operating in increasingly complex environments (Zhang et al., 2023), while industry adoption remains shaped by economic, technical, organizational, and work-culture barriers (Davila Delgado et al., 2019). As robotic systems move from standalone automation tools toward interactive and adaptive systems, construction work is increasingly shifting toward human-robot collaboration (HRC), where workers and robotic systems must coordinate tasks, exchange information, and maintain safety under dynamic site conditions (Zhang et al., 2023).

This shift has important implications for construction engineering and management because the value of robotic technologies depends not only on the capabilities of the robotic systems, but also on the readiness of the workforce that must select, deploy, supervise, operate, maintain, and evaluate them (Bademosi & Issa, 2022). Conventional construction education and workforce training remain largely oriented toward traditional equipment operation, trade-based skill acquisition, and general construction management competencies (Pan et al., 2021). These approaches provide limited preparation for HRC, which requires technical literacy, situational awareness, interface use, task coordination, safety judgment, data interpretation, systems reasoning, and deployment-oriented decision-making (Wu et al., 2024). As a result, construction organizations may invest in robotic technologies without having systematic approaches for identifying the competencies workers need, translating those competencies into training content, or aligning workforce development with technology adoption (Davila Delgado et al., 2019). This creates a management problem as much as an educational problem: robotics adoption can be slowed when workforce capabilities, training pathways, and organizational implementation strategies are not developed alongside the technology (Bademosi & Issa, 2022). Thus, HRC workforce preparation is directly connected to technology and innovation management, safety governance, project delivery, role definition, and organizational readiness for robotic implementation. Recent construction robotics research has increasingly examined HRC competencies, workforce readiness, and robotic applications in the built environment. Prior studies have identified knowledge, skills, and abilities (KSAs) related to robot operation, task planning, sensing, programming, human-robot interfaces, communication, safety, and decision-making (Olukanni et al., 2026b). These studies provide an important foundation by clarifying what construction learners and practitioners need to know and be able to do in HRC-enabled work environments. However, identifying competencies is only the first step. For competencies to support workforce development, they must be translated into cognitively aligned learning objectives (LOs), organized into coherent instructional modules, and validated by experts for relevance, conceptual fit, cognitive appropriateness, and instructional usability.

Parallel bodies of literature provide useful tools for this translation process. Instructional design frameworks such as ADDIE (Analysis, Design, Development, Implementation, and Evaluation) provide a structured process for moving from needs analysis to design, development, implementation, and evaluation (Branch & Varank, 2009; Reiser & Dempsey, 2012). Competency-based education (CBE) emphasizes demonstrable performance outcomes and the alignment of competencies, learning activities, and assessments (Frank et al., 2010). RBT provides a cognitive framework for classifying learning outcomes across knowledge and cognitive process levels, supporting the development of LOs that vary from conceptual understanding to application, analysis, evaluation, and creation (Anderson & Krathwohl, 2001; Krathwohl, 2002). LO design approaches, including the Audience, Behavior, Condition, and Degree (ABCD) structure, further support the formulation of measurable objectives that specify what learners should be able to demonstrate, under what conditions, and to what degree of performance (Mager & Peatt, 1962). These frameworks offer a strong instructional foundation for transforming HRC competencies into training modules.

Furthermore, advances in AI and natural language processing (NLP) provide opportunities to make competency augmentation and curriculum design more scalable and evidence-informed. NLP methods can extract competency-related statements from unstructured sources such as job advertisements, case studies, technical documents, and workforce materials. Sentence-embedding approaches can support semantic grouping, similarity analysis, and

mapping between extracted evidence and competency frameworks. Sentence-embedding models such as Sentence-BERT can generate sentence-level representations that support semantic similarity and clustering tasks (Chai & Chen, 2025). For construction robotics training, these computational methods are valuable because industry-facing sources could reveal emerging workforce demands that may not yet be fully represented in literature-derived competency frameworks.

However, computational methods alone cannot determine whether extracted competencies are instructionally meaningful, practically relevant, or appropriate for construction workforce development. Human review and expert validation remain necessary to ensure that computational outputs are aligned with domain needs and training purposes (Kumar et al., 2024). Despite these advances, existing research remains fragmented. HRC competency studies often stop at competency identification. Instructional design studies provide general models but rarely address the specific demands of construction robotics and collaborative work systems. AI- and NLP-assisted curriculum studies demonstrate the potential of computational support, but remain underdeveloped in construction robotics education (Barasa et al., 2026). The gap, therefore, is that the field lacks a transparent, reproducible, and expert-validated process for augmenting HRC competencies, translating them into LOs, organizing them into theory-constrained training modules, and validating those modules for construction workforce preparation. This gap limits educators, training providers, and construction organizations' ability to translate emerging robotics requirements into actionable workforce development programs.

Addressing these gaps requires an integrated methodological approach that connects competency identification, LO formulation, instructional design, and expert validation into a single, transparent, and reproducible workflow. It also requires a framework to ensure that translated instructional content remains theoretically grounded, contextually relevant to construction practice, and adaptable to evolving robotic applications. Therefore, this study aims to develop and validate training modules for senior undergraduate and graduate students in construction-related disciplines and construction professionals, including early-career construction engineers, field engineers, superintendents, project managers, safety professionals, BIM/VDC personnel, and other practitioners, who are expected to plan, supervise, deploy, operate, maintain, or collaborate with robotic systems during construction project delivery. The proposed NLP-driven methodological framework systematically integrates competency identification, LO formulation, training module development, and expert validation into a unified, transparent, and reproducible workflow. To achieve this aim, the study pursues the following specific objectives:

- To augment an existing HRC competency framework using document-based source extraction, rule-based KSA classification, semantic clustering, and local LLM-assisted candidate mapping to identify evidence-supported competency extensions.
- To develop theory-constrained HRC training modules by integrating ADDIE, CBE, RBT-based cognitive classification, ABCD-informed LO generation, construction robotics contextual grounding, and module-coherence checks.
- To expert-validate the developed training modules for HRC in construction through a Delphi study evaluating competency relevance, module-competency fit, cognitive classification, LO quality, and module completeness.

The study adopts an integrated computational workflow to augment HRC competencies from document-based sources through evidence extraction, KSA classification, semantic clustering, and local LLM-assisted candidate mapping. The augmented competencies were cognitively profiled using RBT, translated into ABCD-informed LOs, contextualized using construction robotics embeddings, organized into theory-constrained modules, and validated through a Delphi study with academic, learning-science, and industry experts. The computational workflow supported evidence extraction, competency augmentation, cognitive classification, LO synthesis, module formation, module-coherence checking, and coverage-matrix generation; final interpretation and validation relied on manual screening, independent reviewer assessment, and Delphi consensus.

This study contributes to knowledge in three ways. First, it provides a reproducible process for augmenting and translating HRC competencies into cognitively aligned learning objectives and training modules. Second, it demonstrates how rule-based extraction, embedding-supported semantic clustering, local LLM-assisted candidate mapping, RBT-based cognitive profiling, ABCD-informed learning objective (LO) generation, and theory-constrained module formation can support transparent instructional design for construction robotics workforce training. Third, it provides expert-validated evidence on the relevance, conceptual fit, cognitive alignment, and LO quality of HRC training modules, demonstrating that assessment requirements differ systematically across

knowledge, skill, and ability competencies and that higher-order ability competencies require more authentic and context-sensitive assessment approaches. These contributions address management and leadership concerns in construction engineering practice by linking robotics adoption to workforce planning, continuing education, training standardization, assessment design, technology implementation, safety governance, and organizational readiness. The remainder of the paper is organized as follows. The next section reviews literature on HRC in construction, workforce competency modeling, instructional design, and AI/NLP-assisted curriculum development. The methodology section presents the integrated workflow for competency augmentation, module development, and Delphi validation. The results section reports the augmented competency framework, cognitive classification, learning objective generation, training module structure, Delphi consensus outcomes, and qualitative feedback. The discussion interprets the findings and their implications for construction education, workforce development, engineering management, and robotics adoption. The paper concludes with contributions, limitations, and directions for future research.

2. BACKGROUND

2.1 Human-robot collaboration in construction

Early construction robotics primarily emphasized mechanization and isolated task automation for repetitive, hazardous, or physically demanding activities such as material handling, surface finishing, bricklaying, inspection, and prefabrication. These systems were often deployed as standalone automation units with limited integration into broader construction workflows. More recent developments reflect a shift toward HRC, where robotic systems function within dynamic socio-technical construction environments rather than as isolated machines (Liang et al., 2021). In this context, robots may serve as task executors, sensing platforms, collaborative tools, mobile agents, assistive devices, or data-generating systems embedded within construction workflows (Liang et al., 2021). T. Bock, & Linner, T. (2015) describes the broader trajectory of construction automation toward increasingly capable robotic systems, while adoption studies show that implementation remains constrained by economic, technical, cultural, and organizational barriers (Davila Delgado et al., 2019). HRC in construction includes supervisory control, shared autonomy, teleoperation, human-supervised robotic task execution, and physical collaboration (Liang et al., 2021). Across these interaction forms, humans may provide intent, sequencing, task oversight, safety judgment, and intervention, while robots perform sensing, movement, manipulation, inspection, or data collection. Effective HRC, therefore, depends not only on robotic autonomy but also on the human collaborator's ability to understand system capabilities, interpret robotic behavior, anticipate risks, and intervene appropriately under uncertainty (Zhang et al., 2023). Existing HRC research establishes that construction robotics is not only a technical implementation issue but also a workforce capability issue (Ojha et al., 2025). Construction environments are dynamic, unstructured, and safety-critical (Liang et al., 2021), making HRC training more complex than teaching learners how to operate a new tool. Workers and managers must be prepared to evaluate robotic systems, coordinate human and robotic work, manage safety and productivity implications, and assess whether robotic deployment creates operational value. However, existing construction education and workforce training programs have not yet fully incorporated the interdisciplinary competencies required for collaborative robotic work, including robotics literacy, human-robot interface use, sensor interpretation, robot task planning, data analytics, safety reasoning, system integration, and technology adoption decision-making.

2.2 Workforce competency modeling for construction robotics

The growing interest in construction robotics has prompted increased attention to the workforce competencies required for effective HRC. Competency frameworks are useful because they define the KSAs that learners and practitioners need to perform in a domain. In construction robotics, prior studies have identified competencies related to robot types, construction robot applications, sensing, programming, control systems, human-robot interfaces, safety standards, task planning, communication, teamwork, decision-making, and problem-solving (Olukanni et al., 2026b). These studies clarify the multidimensional nature of HRC, which requires technical knowledge, procedural skills, cognitive abilities, social coordination, and safety judgment (Bademosi & Issa, 2022). However, competency identification by itself is insufficient for workforce training. A list of competencies does not indicate how those competencies should be sequenced, converted into LOs, grouped into modules, or assessed. This distinction is critical for construction robotics education because competencies such as sensing, programming, safety management, robot operation, communication, systems thinking, and deployment evaluation

do not develop independently. They must be organized into learning pathways that allow learners to progress from foundational concepts to applied and higher-order capabilities (Enke et al., 2015). CBE provides a useful foundation for this translation because it emphasizes demonstrable outcomes and alignment among competencies, instruction, and assessment. Frank et al. (2010) describe CBE as an outcomes-focused approach organized around predefined abilities and accountable learner performance. For HRC training, this orientation shifts attention from technology exposure alone to what learners should be able to do with, around, and in relation to robotic systems. Nevertheless, CBE does not automatically determine the cognitive level, module structure, or assessment strategy for each competency. Prior studies on construction robotics and automation often identify competency needs or adoption barriers but provide limited guidance on how to translate competencies into coherent instructional modules. This creates a gap between competency modeling and instructional implementation.

2.3 Instructional design frameworks and learning taxonomies

Instructional design frameworks provide systematic procedures for translating learning needs into educational interventions. ADDIE remains widely used because it provides a clear structure for analysis, design, development, implementation, and evaluation (Branch & Varank, 2009). In workforce-oriented and technical training contexts, ADDIE supports traceability from needs analysis to instructional design outputs and evaluation (Branch & Varank, 2009). For HRC training, ADDIE provides a process logic for moving from competency augmentation to LO development, module formation, and expert validation. CBE complements ADDIE by emphasizing performance-based outcomes and alignment between competencies, learning activities, and assessment (Henri et al., 2017). However, CBE requires additional cognitive structuring to determine the expected level of understanding, application, analysis, evaluation, or creation. RBT provides this cognitive structure by distinguishing among knowledge dimensions (KD)—factual, conceptual, procedural, and metacognitive—and cognitive process dimensions (CPD), including remember, understand, apply, analyze, evaluate, and create (Anderson & Krathwohl, 2001; Krathwohl, 2002). For HRC training, RBT is useful because the competency set spans multiple levels of cognitive demand: learners must understand robotic systems, apply operational procedures, analyze sensor and performance data, evaluate safety and deployment decisions, and create or adapt system integration strategies. LOs further require explicit performance statements. Mager's seminal work on instructional objectives emphasizes the need to specify observable performance, conditions, and criteria (Mager & Peatt, 1962). The Audience, Behavior, Condition, and Degree (ABCD) structure defines this logic by specifying who performs the action, what behavior is expected, under what conditions, and to what degree. This is relevant for HRC because LOs must be grounded in construction robotics contexts rather than written as generic statements. ADDIE, CBE, RBT, and ABCD provide the instructional foundation for converting HRC competencies into assessable training modules. However, these frameworks have rarely been integrated with computational competency augmentation and Delphi validation in construction robotics training.

2.4 Computationally supported curriculum design

Computational methods are increasingly being used to support curriculum analytics, skill extraction, competency mapping, LO alignment, and training design. In the workforce and education domains, unstructured and semi-structured documents can contain evidence about emerging skill demands, role expectations, technology requirements, and instructional gaps (Saleema et al., 2025). For construction robotics, such sources are useful because HRC competencies span technical knowledge, procedural skill, safety reasoning, communication, systems thinking, and deployment-oriented judgment. Sentence-embedding methods are particularly useful for semantic comparison because they represent text in a way that allows similarity analysis across competency statements, LOs, module descriptions, and robot-application contexts (Chai & Chen, 2025). Sentence-BERT produces sentence-level embeddings that can be compared using cosine similarity, making it suitable for semantic search, contextual matching, semantic clustering, and coherence checking (Reimers & Gurevych, 2019). In HRC training design, such methods can support the alignment of competencies with relevant construction robotics scenarios and help assess whether competencies assigned to a module are semantically coherent. Large language models can also support competency mapping by comparing extracted evidence with an existing competency framework and suggesting provisional mappings or candidate extensions (Romão et al., 2025). However, LLM outputs can be sensitive to prompt design and may overgenerate categories. Therefore, LLM-assisted curriculum development should be implemented as a human-in-the-loop process in which model outputs are screened, interpreted, and validated by researchers and domain experts. For HRC training, computational methods can improve traceability

and scalability, but they do not replace expert judgment about relevance, conceptual fit, instructional usability, or completeness.

2.5 Research gaps

The reviewed literature reveals three interconnected gaps. First, construction robotics and HRC studies have advanced understanding of robotic applications, interaction requirements, and workforce competencies, but they provide limited guidance on how competencies should be translated into training modules. Second, instructional design and learning-taxonomy research provide strong foundations through ADDIE, CBE, RBT, and ABCD-aligned LO design, but these frameworks have not been sufficiently integrated into construction robotics and HRC training. Third, computational methods offer scalable tools for evidence extraction, semantic comparison, competency mapping, and instructional design support, but their use in construction-focused HRC curriculum development remains limited. These gaps indicate that the field lacks an integrated, theory-guided, computationally supported, and expert-validated methodology for converting HRC competencies into training modules for construction workforce development. Existing approaches remain fragmented across competency identification, instructional theory, AI-assisted extraction, and expert validation. As a result, construction education programs and industry training providers have limited guidance for standardizing HRC training, aligning workforce competencies with robotic system deployment, or validating whether proposed modules are relevant, cognitively appropriate, and instructionally usable. This study addresses these gaps by integrating document-based source extraction, rule-based KSA classification, Sentence-BERT semantic clustering, local LLM-assisted candidate mapping, ADDIE-guided instructional design, CBE, RBT-based cognitive classification, ABCD-informed LO generation, theory-constrained module formation, module-coherence checking, coverage-matrix generation, and Delphi-based expert validation into a unified workflow.

3. RESEARCH FRAMEWORK AND METHODOLOGY

3.1 Overview of research framework

This study adopted an instructional design and expert-validation framework to develop HRC training modules for construction learners and professionals who are expected to plan, supervise, manage, or collaborate with robotic systems during construction project delivery. The design is guided by the ADDIE model, which supports systematic translation of learning needs into instructional outputs and is widely used in instructional design and technical training contexts (Branch & Varank, 2009). The analysis, design, development, and evaluation phases were completed, while the implementation phase was excluded because the focus was on pre-deployment design and expert validation of the modules. In the analysis phase, a baseline HRC competency framework from Olukanni et al. (2026b) was augmented using a computational pipeline that processed Portable Document Format (PDF) sources and extracted competency evidence. The workflow included embedded text extraction, optical character recognition (OCR) fallback, sentence segmentation, HRC-context filtering, rule-based KSA classification, deduplication, Sentence-BERT semantic clustering, internal validation checks, and LLM-assisted generation of candidate mapping to existing or provisional competency codes, which were manually verified against the source documents and predefined competency framework before Delphi validation. In the design phase, the augmented competencies were cognitively profiled using RBT, aligned with CBE principles, and translated into LOs using an ABCD-informed structure. In the development phase, competencies and LOs were organized into theory-constrained HRC training modules. In the evaluation phase, modules were assessed through internal computational validation, module-coherence checks, independent reviewer evaluation of the augmented competency framework, and a two-round Delphi process. The framework presented in Figure 1 integrates four components: (1) computational HRC competency augmentation, (2) RBT-based cognitive profiling and ABCD-informed LO generation, (3) embedding-supported contextual grounding with theory-constrained module design, and (4) multi-layer validation including coherence checks, coverage assessment, independent review, and Delphi validation. This structure ensures full traceability from source evidence to competencies, competencies to LOs, LOs to modules, and modules to expert-validated instructional structures. It also aligns with construction engineering management needs by linking competency identification with training standardization, workforce planning, and construction robotics adoption.

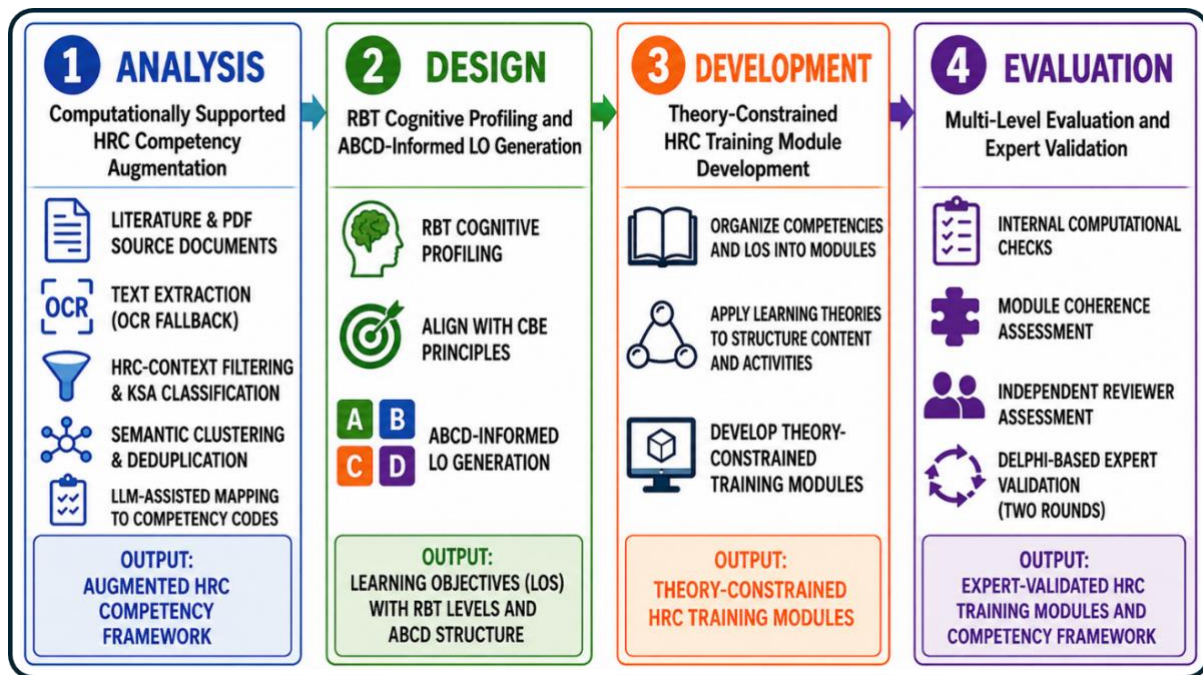


Figure 1: Overall research framework and methodology for HRC training module development and expert validation.

3.2 Computationally supported HRC competency augmentation

The computational workflow developed to augment the baseline HRC competency framework integrates PDF text extraction, HRC-based filtering, rule-based KSA classification, semantic clustering, internal validation diagnostics, and local LLM-assisted competency mapping to systematically identify, validate, and expand HRC competencies in construction. The overall workflow and augmentation process are illustrated in Figure 2.

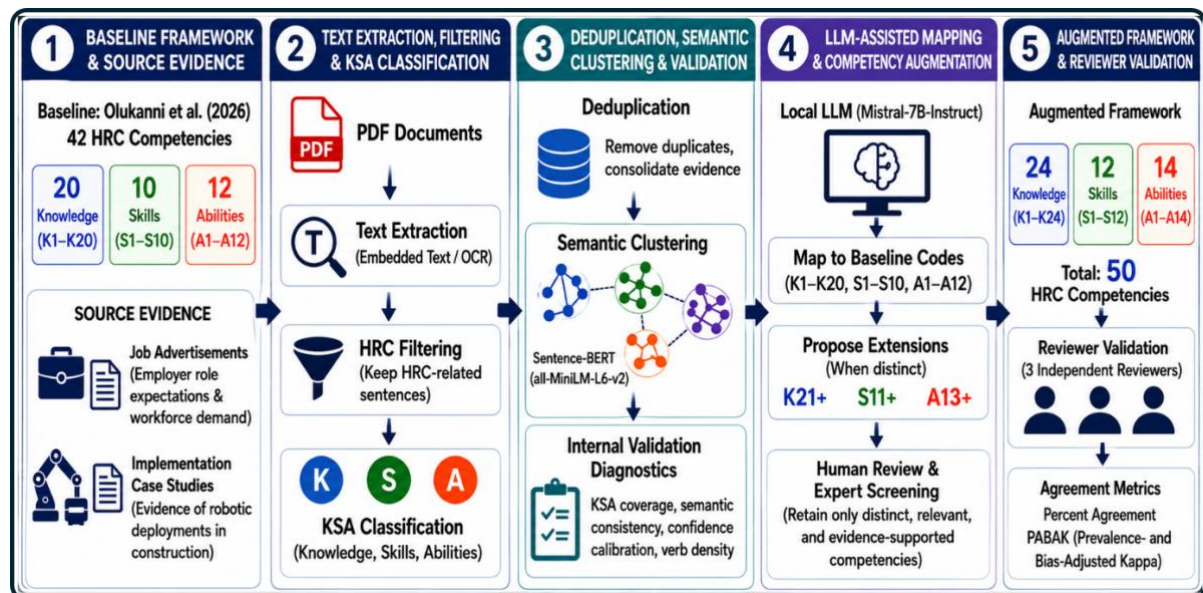


Figure 2: NLP-driven competency augmentation and validation pipeline.

3.2.1 Baseline competency framework and source evidence

The analysis phase focused on identifying and augmenting the competencies required for effective HRC in construction. The baseline competency framework was derived from a prior systematic literature review by Olukanni et al. (2026b) and included 42 HRC competencies, consisting of 20 knowledge areas, 10 skills, and 12 abilities. This baseline framework served as the starting point for this study. Because HRC technologies, workforce roles, and construction robotics applications are evolving rapidly, the baseline framework was augmented using industry-facing documentary evidence. Two types of source evidence were used: construction robotics-related job advertisements and construction industry robotic implementation case studies. Job advertisements were used because they reflect employer-defined role expectations, workforce demand signals, and emerging capability requirements, while implementation case studies were used because they provide evidence of how robotic systems are deployed in construction contexts and the associated operational, technical, safety, and managerial capabilities associated with those deployments (Oh Heung et al., 2023; Zheng et al., 2020). Source documents were retained when they contained HRC- or construction-robotics-related evidence, including references to robots, cobots, human-robot interaction, collaborative robotics, exoskeletons, Unmanned Aerial Vehicles (UAVs), drones, autonomous systems, sensing, inspection, monitoring, mapping, lifting, assembly, transport, robotic task execution, or construction automation. Documents unrelated to construction robotics, HRC, automation, or workforce capability requirements were excluded. The retained corpus provided the input for competency extraction and augmentation.

3.2.2 PDF extraction, HRC filtering, and rule-based KSA classification

The competency extraction process used a multi-stage hybrid computational pipeline to extract, structure, and validate competency-related statements from source documents. Where available, embedded text was extracted directly; otherwise, OCR was applied to convert scanned or image-based PDFs into machine-readable text, consistent with established document-processing approaches (Surana et al., 2022). The resulting text was segmented into sentences, as competency evidence in job advertisements and case studies is typically expressed as brief task or technology descriptions, role expectations, or capability statements. A domain-specific filtering step was then applied using an HRC context gate. Sentences were retained only if they contained HRC or robotics-related terms such as robot, cobot, human-robot, exoskeleton, UAV, drone, autonomous, robot arm, industrial robot, mobile robot, and gantry robot, representing examples from the broader HRC keyword dictionary used in the computational pipeline. The sentence-retention rule was represented in Equation (1) as:

$$I(s_i) = \begin{cases} 1, & \text{if } \text{tokens}(s_i) \cap K_{HRC} \neq \emptyset \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where s_i is an extracted sentence, K_{HRC} is the predefined set of HRC-related signal terms, and $I(s_i) = 1$ indicates that the sentence was retained for KSA classification. The complete keyword dictionary and implementation are available in the publicly archived Virginia Tech Research Data Repository (Olukanni et al., 2026a). Retained sentences were subsequently classified into KSA categories using rule-based linguistic indicators. Knowledge statements were identified through references to concepts such as robotics systems, sensors, control systems, simulation, machine learning, standards, and programming. Skill statements were identified through action-oriented verbs such as operate, program, inspect, monitor, maintain, configure, implement, test, and verify. Ability statements were identified through descriptors related to teamwork, communication, problem-solving, adaptability, decision-making, critical thinking, and safety. Each extracted statement was assigned a KSA category, confidence score, evidence page reference, and contextual indicators such as robot type and application, ensuring an auditable and traceable extraction process prior to competency augmentation.

3.2.3 Deduplication, semantic clustering, and internal validation diagnostics

Following rule-based extraction, duplicate and overlapping statements were consolidated. Deduplication was performed using KSA type and phrase similarity, while preserving associated metadata, including confidence scores, robot-type indicators, application context, and page-level evidence references to maintain full provenance. The remaining statements were transformed into vector representations using the Sentence-BERT model (all-MiniLM-L6-v2), selected for efficient sentence-level semantic similarity modeling compared to standard BERT approaches (Reimers & Gurevych, 2019). Semantic similarity between two statements was computed using cosine similarity, as shown in Equation (2).

$$\text{sim}(s_i, s_j) = \frac{e(s_i) \cdot e(s_j)}{\|e(s_i)\| \|e(s_j)\|} \quad (2)$$

where $e(s_i)$ and $e(s_j)$ are the Sentence-BERT embeddings of statements s_i and s_j , respectively. Hierarchical agglomerative clustering with cosine distance was then applied to group semantically similar statements. Cluster representatives were selected from the highest-confidence statements within each cluster, ensuring that representative evidence was retained while redundancy was reduced. Importantly, clustering was not used to automatically define competencies; rather, it supported semantic aggregation and produced a reduced, reviewable set of evidence statements for mapping. Prior to competency mapping, internal validation diagnostics were conducted. Coverage validation assessed the distribution of extracted statements across KSA categories. Semantic consistency validation examined whether clustered statements predominantly reflected a single KSA category. Confidence calibration evaluated whether confidence scores were driven by evidence frequency or rule-based classification clarity. Structural validity was assessed using verb-density analysis to determine whether skill-related statements exhibited higher concentrations of action-oriented verbs compared to knowledge or ability statements. These diagnostics were used to evaluate extraction quality, redundancy reduction, KSA coverage, semantic coherence, and potential classification issues prior to manual screening and LLM-assisted mapping.

3.2.4 Local LLM-assisted mapping and competency augmentation

After semantic clustering, cluster representatives were mapped to the baseline KSA framework using a locally instruction-tuned LLM. A local Mistral-7B-Instruct model, implemented using llama-cpp-python, was used with deterministic settings, consistent with recent applications of instruction-tuned open-weight LLMs for controlled text classification and mapping tasks (Chae & Davidson, 2026). The model received the baseline HRC competency framework (K1–K20, S1–S10, and A1–A12) and was prompted to map extracted statements to existing KSA codes where possible, while proposing provisional extension codes (e.g., K21+, S11+, A13+) only when evidence indicated competencies not represented in the baseline framework. Specifically, the Mistral-7B-Instruct model was provided with the following instruction:

"You are given (1) the canonical HRC competency framework containing the existing Knowledge (K), Skills (S), and Abilities (A) competency definitions and codes, and (2) a batch of extracted competency statements. For each statement, map it to the most appropriate existing competency code (K, S, or A). Only if a statement cannot be reasonably mapped to any existing competency should you propose a new competency using the next available code (K21+, S11+, or A13+). Do not modify existing competency definitions. Return the assigned competency code, justification for the mapping, and indicate whether a new competency is proposed."

The LLM functioned solely as a candidate competency statement-mapping and extension-identification aid, not as an autonomous validator or final decision-maker, and all outputs were constrained by the predefined HRC competency framework. All generated mappings and proposed extensions were reviewed against the extracted statements, source evidence, KSA definitions, and baseline framework. Extensions were retained only when they were semantically distinct, instructionally meaningful, relevant to construction HRC, and supported by evidence. Candidate mappings that lacked documentary evidence, duplicated existing competencies, used overly broad language, exhibited inconsistent interpretations, lacked relevance to construction robotics, or reflected overgeneration were removed or merged with existing competency labels. Final competency inclusion, naming, and KSA grouping were determined through manual screening and expert review. The augmentation process expanded the baseline framework from 42 to 50 HRC competencies, including 24 HRC knowledge areas, 12 skills, and 14 abilities. The resulting framework served as the stable input for cognitive classification, LO generation, module development, and Delphi validation. Following manual screening, the augmented framework was evaluated by three independent domain reviewers separate from the Delphi panel. Reviewers assessed classification accuracy, aggregation decisions, competency labeling, and KSA grouping. Percent agreement was calculated to measure direct reviewer agreement, while prevalence-adjusted and bias-adjusted kappa (PABAK) was calculated because reviewer decisions were likely skewed toward agreement after initial screening. PABAK provides a prevalence- and bias-adjusted reliability estimate under such conditions (Byrt et al., 1993). This reviewer-validation stage assessed the reliability of competency augmentation and labeling, whereas the subsequent Delphi study evaluated the developed training modules. The next stage translated the augmented

competency framework into instructional design outputs, including cognitive profiles, LOs, module assignments, coverage outputs, and coherence outputs.

3.3 Theory-constrained instructional module development

This section presents the theory-constrained instructional module development process that integrates RBT, competency-based instructional design principles, contextual grounding in construction robotics applications, and theory-guided module organization to ensure instructional coherence, cognitive alignment, and traceability across the training framework. As illustrated in Figure 3, the workflow consisted of four interconnected stages: cognitive demand profiling, LO generation and contextual grounding, theory-constrained module formation, and module-coherence and coverage checks. The module architecture assumes learners possess general construction knowledge but not advanced expertise in robotics. Consequently, competencies progress from foundational robotics concepts toward collaborative operation, systems thinking, deployment evaluation, and organizational implementation, allowing the framework to accommodate both university instruction and professional workforce development programs. As detailed in the sections below, the developed instructional outputs were first subjected to an internal computational evaluation to assess structural consistency before human review. This evaluation verified Bloom classification completeness, module assignment consistency, learning objective generation, competency coverage, and within-module semantic coherence using embedding-based similarity measures. This is then followed by Delphi evaluation of the training modules developed.

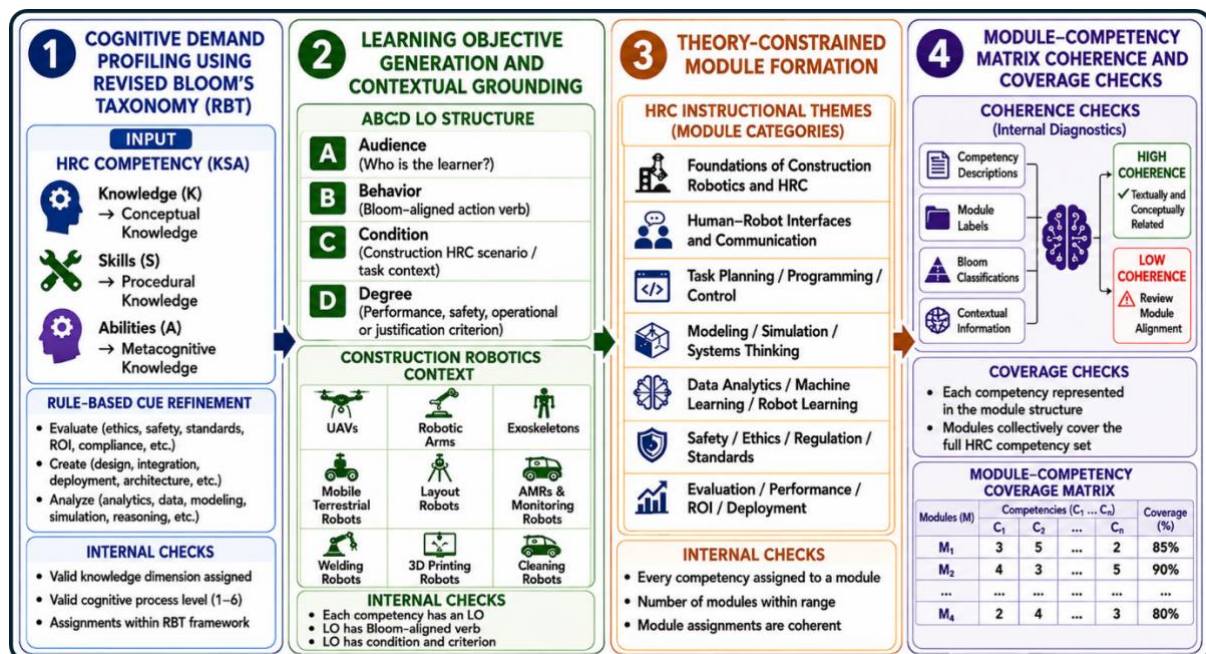


Figure 3: Theory-constrained HRC instructional module development workflow.

3.3.1 Cognitive demand profiling using revised Bloom's taxonomy

Each HRC competency was cognitively profiled using a hybrid rule-based procedure grounded in RBT. KD were inferred from KSA type: knowledge competencies were assigned to factual/conceptual knowledge, skills to procedural knowledge, and abilities to metacognitive knowledge. The initial mapping of KSA types to KD was informed by competency-based education literature and the RBT, which distinguishes declarative knowledge, procedural performance, and self-regulatory capabilities as progressively different forms of learning (Anderson & Krathwohl, 2001; Krathwohl, 2002; Mulder, 2014). Accordingly, knowledge competencies were initially associated with factual and conceptual knowledge, skills with procedural knowledge, and abilities with metacognitive knowledge as a computational starting heuristic rather than a fixed classification. Cognitive process levels were inferred using KSA-type defaults and domain-specific cue rules, consistent with RBT's distinction between KD and CPD (Anderson & Krathwohl, 2001; Krathwohl, 2002). This hybrid approach was necessary because many HRC competencies are expressed as noun phrases, such as “sensors,” “robot control system,”

“systems thinking,” or “robot operation and maintenance,” rather than as explicit action statements. The initial heuristic was subsequently refined through domain-specific cue rules and later evaluated through independent reviewer assessment and Delphi validation, allowing preliminary assignments to be revised where the competency statements reflected different cognitive demands. The rule-based procedure used type-based defaults as a conservative starting point. HRC knowledge areas were initially associated with conceptual understanding, skills with procedural application, and abilities with analytical or metacognitive reasoning. Domain-specific cue rules then refined these assignments. Cues related to ethics, regulation, standards, compliance, safety, cost, return on investment (ROI), and evaluation were used to identify competencies requiring evaluative reasoning. Cues related to integration, design, architecture, and deployment were used to identify competencies requiring creative or design-oriented reasoning. Cues related to analytics, data, machine learning, modeling, simulation, reasoning, and critical thinking were used to identify analytical demand.

Internal checks verified that all competencies received valid KD and CPD and that the resulting categories fell within the RBT framework. These checks supported transparency and consistency, but they were not treated as evidence of instructional effectiveness. Instead, they ensured that the cognitive profiling procedure was internally coherent before translating the competencies into LOs and modules.

3.3.2 Learning objective generation and contextual grounding

Each competency was translated into an LO using an ABCD-informed structure and Bloom-aligned action verbs. The Audience was the learner; the Behavior was selected from a Bloom-aligned verb set; the Condition described a construction HRC scenario or task context; and the Degree specified a performance, safety, operational, or justification criterion, consistent with LO writing principles emphasizing observable performance, conditions, and criteria (Mager & Peatt, 1962). The computational workflow generated LOs by combining competency descriptions, Bloom-aligned verbs, and construction robotics contexts. To ground the objectives in construction robotics practice, a robot-application context dataset was developed. This dataset linked construction robot types to representative applications, including UAVs for monitoring, surveying, progress tracking, inspection, and mapping; robotic arms for drywall installation, tile laying, sheathing, timber framing, and waste sorting; exoskeletons for material handling and lifting support; mobile terrestrial robots for ground-based monitoring and material transport; autonomous layout robots for layout mapping; autonomous mobile robots for safety monitoring and reality capture; welding robots for joining; 3D printing robots for additive construction; and construction cleaning robots for site cleaning. Robot and application contexts were represented as text and compared with competency descriptions using Sentence-BERT embeddings. The workflow selected the most semantically relevant robot-application contexts for each competency and incorporated these contexts into the generated LOs. This step ensured that the LOs were not generic statements about robotics but were linked to plausible construction applications.

Internal validation checks verified that each competency had a corresponding LO, that each LO contained the expected Bloom-aligned action verb, and that each objective included a condition and performance criterion. These checks supported traceability and instructional completeness, but the generated LOs were still treated as design artifacts requiring expert review. This distinction is important because computationally generated LOs can satisfy formal structure while still requiring refinement for clarity, measurability, and construction-specific assessment design.

3.3.3 Theory-constrained module formation

Training modules were formed using a theory-constrained assignment strategy. The modules were formed using predefined HRC-relevant instructional themes and then checked for semantic coherence. This design choice was used because unsupervised clusters may produce statistically coherent groupings that are difficult to interpret instructionally, especially when HRC competencies span technical, cognitive, safety, social, and organizational domains. The module-formation workflow assigned competencies to modules using theme-based cue rules. The thematic categories reflected major instructional domains in HRC training, including foundations of construction robotics and HRC, human-robot interfaces and communication, task planning, programming, and control, modeling, simulation, and systems thinking, data analytics, machine learning and robot learning, safety, ethics, regulation and standards, and evaluation, performance, ROI, and deployment. Where needed, semantic similarity was used as a fallback to support assignment decisions, but it did not define the module structure. This process

reflected a constraint-first design logic: module boundaries were guided by instructional coherence, Bloom progression, task relevance, and cognitive-load considerations rather than by clustering alone.

Internal validation checks verified that every competency was assigned to a module, that the number of modules remained within a manageable instructional range, and that module assignments were semantically coherent. A coverage matrix was constructed to represent Bloom-based cognitive depth associated with each competency within the module architecture. The values in the coverage matrix represented instructional depth, not competency importance or frequency. Higher values indicated that a competency was addressed at a higher cognitive level within the module structure. This matrix provided a transparent bridge between cognitive classification, LO generation, module structure, coverage assessment, and expert validation.

3.3.4 Module-competency coherence and coverage checks

After the module assignment, module coherence and coverage checks were conducted to evaluate whether the training architecture was internally consistent. Semantic coherence was assessed by comparing embedded representations of competency descriptions, module labels, Bloom classifications, and contextual information. These checks were used as internal diagnostics. A high semantic coherence score indicated that competencies assigned to a module were textually and conceptually related, while a lower score signaled that a module might require closer review. Coverage checks verified that each competency was represented in the module structure and that the modules collectively covered the full HRC competency set. The coverage matrix linked competencies to modules and represented the cognitive depth at which each competency was addressed. This output made the module design process auditable by showing how the augmented competency framework was translated into instructional units and cognitive expectations.

3.4 Delphi-Based Expert Validation

3.4.1 Delphi Validation Design

The developed HRC training modules were validated through a two-round Delphi study. The Delphi method was selected because it supports consensus-building among experts and is appropriate for validating frameworks in emerging domains where direct empirical implementation evidence is still limited (Hasson & Keeney, 2011). In this study, the Delphi process served as the expert-evaluation phase of the ADDIE-guided framework. The computational workflow generated instructional design artifacts, including augmented competencies, cognitive profiles, LOs, module assignments, coverage outputs, and coherence outputs. However, computational coherence does not by itself establish instructional relevance, conceptual fit, or practical usability. Expert consensus was therefore used to evaluate whether the proposed HRC training modules were relevant, conceptually coherent, cognitively appropriate, instructionally usable, and sufficiently complete before future classroom or industry implementation.

3.4.2 Questionnaire development and pilot testing

The Delphi questionnaire was developed from the module architecture, competency assignments, RBT cognitive classifications, and LOs generated in earlier phases. The instrument evaluated five dimensions: competency relevance to the assigned module, module-competency fit, cognitive classification, LO quality, and module completeness. LO quality was assessed across clarity, measurability, and usability. The questionnaire was implemented in Qualtrics and underwent two stages of pilot testing before distribution to the Delphi panel. First, 10 participants with backgrounds aligned with construction robotics, HRC, or workforce training reviewed the questionnaire for clarity, survey flow, response burden, and usability. Second, cognitive interview piloting was conducted with 5 participants to examine whether they interpreted questions and response options as intended. Feedback from pilot testing led to refinements to item wording, response scale presentation, module descriptions, question sequencing, and alignment among competencies, cognitive classifications, and LOs. Pilot testing was included to reduce measurement error before the Delphi rounds. The survey asked experts to evaluate complex instructional artifacts, including competencies, cognitive classifications, module assignments, and LOs; therefore, it was important to verify that the instrument was understandable and that participants interpreted the questions as intended.

3.4.3 Expert panel selection

A panel of 15 experts was purposively selected for the Delphi study. The panel included 9 construction robotics professors, 3 learning sciences or engineering education professors, and 3 industry professionals with experience in construction robotics or digital construction practice in the United States. The Delphi panel was distinct from the participants involved in pilot testing and from the reviewers involved in competency screening and interpretation. Experts were selected based on demonstrated expertise in HRC, construction robotics, learning sciences, instructional design for emerging technologies, or workforce training. Eligibility was demonstrated through peer-reviewed publications, funded research, industry project involvement, or at least 5 years of relevant professional experience. Participants were identified through a targeted review of relevant literature, professional networks, and conference participation records and were invited by email. Upon agreement to participate, each expert received study information and consent materials approved by the Virginia Tech Institutional Review Board. The panel composition was designed to balance three forms of expertise: construction robotics domain knowledge, instructional and learning-science expertise, and industry implementation experience. The predominance of academic experts was intentional because the Delphi evaluated competency structure, cognitive classification, learning objective quality, module organization, and instructional coherence, in addition to technical relevance. These aspects require expertise in competency-based education, instructional design, and construction robotics research, while the inclusion of experienced industry practitioners ensured that the framework remained grounded in contemporary HRC implementation and workforce needs. This balance was necessary because the study evaluated not only whether the competencies were technically relevant, but also whether they were instructionally coherent and practically meaningful for workforce preparation. All 15 invited experts agreed to participate in the Delphi study and completed all two rounds, resulting in a 100% response rate and 100% participant retention throughout the study. No panel attrition occurred between rounds, ensuring consistency in expert feedback and consensus development across the iterative evaluation process.

3.4.4 Delphi rounds and consensus measurement

Round 1 focused on validating the internal structure of the HRC training modules. Experts evaluated the relevance of each competency to its assigned module using a 5-point Likert scale ranging from 1 = Very low to 5 = Very high. Module–competency fit was evaluated using a binary Yes/No judgment indicating whether each competency was appropriately assigned to its module. For “No” responses, participants were asked to suggest an alternative module. Experts also evaluated the cognitive classification of each competency by assessing whether the assigned KD and CPD of RBT were appropriate. For “No” responses, participants were asked to suggest alternative classifications. LO quality was evaluated across clarity, measurability, and usability using 5-point Likert scales. Experts provided qualitative explanations for disagreement, suggested refinements, and commented on module completeness.

Round 1 quantitative ratings and qualitative comments were synthesized to identify items requiring revision. Items that did not meet the consensus threshold or received substantive qualitative concerns were refined. The revised materials were redistributed in Round 2, where experts reassessed non-consensus or revised items and indicated whether the updated modules adequately reflected the feedback. Round 2 was therefore used to determine whether expert agreement improved after structured feedback and revision and to identify any persistent areas of disagreement. For Likert-scale items, consensus was defined as both a median rating of at least 4 and at least 70% agreement, with agreement defined as the percentage of experts rating an item 4 or 5. For binary Yes/No items, consensus was defined as at least 70% “Yes” responses. These criteria were applied consistently across competency relevance, module–competency fit, cognitive classification, LO quality, and module completeness, consistent with established Delphi practice, where a 70% threshold is widely accepted as representing meaningful expert consensus while accommodating disciplinary diversity (Diamond et al., 2014; Jünger et al., 2017). For the present panel of 15 experts, the nominal 70% threshold required agreement from at least 11 participants, as 10 of 15 experts corresponds to 66.7%, whereas 11 of 15 corresponds to 73.3%. Thus, an item could achieve consensus while up to four experts expressed disagreement, reflecting a pragmatic level of consensus appropriate for a heterogeneous interdisciplinary expert panel rather than universal agreement.

3.5 Data analysis

Data analysis included separate procedures for computational workflow checks, independent reviewer assessment, and Delphi consensus analysis. The computational workflow assessed extraction outputs, mapping support,

internal consistency, traceability, module coverage, and semantic coherence, while the Delphi study assessed whether the developed training modules were relevant, coherent, cognitively appropriate, instructionally usable, and sufficiently complete from the perspective of domain experts. For the competency augmentation stage, agreement among the three independent reviewers was assessed using percent agreement and PABAK. PABAK was used because agreement decisions were expected to be skewed toward agreement after initial screening, and PABAK provides a prevalence- and bias-adjusted reliability estimate under such conditions (Byrt et al., 1993). LLM-generated mappings were reviewed as candidate outputs rather than final decisions. The final augmented competency framework reflected source evidence, semantic distinctiveness, manual screening, and expert interpretation. For the module-development stage, outputs were examined to verify that all competencies received valid knowledge-dimension and cognitive-process assignments, that all competencies were linked to LOs, that each competency was assigned to a module, and that module assignments were semantically coherent. Coverage outputs were used to examine the relationship between modules, competencies, and cognitive depth. For Delphi analysis, quantitative responses were analyzed using descriptive statistics. For Likert-scale items, consensus was defined a priori as a median rating of at least 4 and at least 70% agreement, where agreement was operationalized as a rating of 4 or 5. For binary Yes/No items, consensus was defined as at least 70% Yes responses. These criteria were applied to competency relevance, module–competency fit, cognitive classification, LO quality, and module completeness. Qualitative feedback from Delphi experts was analyzed thematically. Comments were iteratively coded to identify recurring concerns related to module boundaries, competency placement, RBT classification, LO clarity, LO measurability, usability, and module completeness. Quantitative results and qualitative comments were interpreted together to guide revisions between Delphi rounds and to explain areas of persistent disagreement. Quantitative analyses were conducted using Microsoft Excel and Jamovi, while qualitative comments were manually coded and organized in Microsoft Excel because the volume of qualitative feedback was manageable.

4. RESULTS

4.1 HRC competency augmentation and validation

4.1.1 Augmentation of HRC competencies

The competency augmentation process extended the baseline HRC competency framework reported by Olukanni et al. (2026b), which included 42 competencies (20 knowledge, 10 skills, and 12 abilities) (see Table 1). Consistent with the methodology, the augmentation process used an NLP-assisted, embedding-supported, and LLM-assisted workflow to extract competency evidence from construction robotics job advertisements and robotic implementation case studies, map the extracted evidence against the baseline KSA framework, and identify distinct competency extensions. Evidence extracted from industry job advertisements and case studies of robotic implementation in construction yielded 490 competency-related statements. After duplicate removal, 463 statements remained, including 419 knowledge-related statements, 36 skill-related statements, and 8 ability-related statements. Manual screening removed incomplete, unclear, or contextually irrelevant statements, resulting in 418 screened and retained statements, including 380 knowledge-related statements, 30 skill-related statements, and 8 ability-related statements. These retained statements were semantically aggregated into clusters, and cluster representatives were mapped to the baseline competency framework using the local LLM-assisted mapping procedure. The mapping identified statements that aligned with existing K1–K20, S1–S10, and A1–A12 codes, as well as candidate extension codes that were subsequently screened and reviewed.

Table 1: Baseline HRC competencies identified from literature by Olukanni et al. (2026b).

Code	HRC knowledge	Code	HRC skills	Code	HRC abilities
K1	Types of robots	S1	Effective communication	A1	Teamwork
K2	Construction robot applications	S2	Task planning	A2	Communication
K3	Robot anatomy and technical specifications	S3	Regulation standards compliance	A3	Continuous learning
K4	Sensors	S4	Safety management	A4	Problem-solving
K5	Task planning	S5	Technical skill	A5	Adaptability
K6	HRC ethics and regulation	S6	Robot programming	A6	Attention to detail
K7	HRC safety and standards	S7	Data analytics and management	A7	Analytical aptitude



Code	HRC knowledge	Code	HRC skills	Code	HRC abilities
K8	HRC evaluation	S8	Human-robot interface proficiency	A8	Decision-making
K9	HRC-related fields	S9	Application of machine learning algorithms	A9	Critical thinking
K10	Immersive virtual environments	S10	Simulation and modeling	A10	Spatial awareness
K11	Communication modes and technologies			A11	Cultural and social awareness
K12	Human-robot interface			A12	Safety awareness
K13	Robot control system				
K14	System integration				
K15	Programming				
K16	Modeling and simulation				
K17	Data analytics and machine learning				
K18	Robot learning methods				
K19	Computational design				
K20	Robot operating system (ROS)				

The augmentation process identified eight evidence-supported competency extensions not explicitly represented in the baseline framework. The newly identified knowledge areas were cost and ROI in robotics, robotic project management, robotics sustainability, and robotics-related jobs. The newly identified skills were robot operation and maintenance and systems thinking. The newly identified abilities were robotic system reasoning and robot acceptance and operational reliability. These additions expanded the framework from 42 to 50 competencies, including 24 knowledge areas, 12 skills, and 14 abilities (Table 2).

Table 2: HRC competency augmentation summary.

Category	Initial HRC competencies	Augmented HRC competencies	Newly identified HRC competencies
Knowledge	20	24	K21 – Cost and ROI in robotics
			K22 – Robotic project management
			K23 – Robotics sustainability
			K24 – Robotics-related jobs
Skills	10	12	S11 – Robot operation and maintenance
			S12 – Systems thinking
Abilities	12	14	A13 – Robotic system reasoning
			A14 – Robot acceptance and operational reliability

4.1.2 Internal and external validation of the augmented HRC competency framework

Internal validation examined construct coverage, semantic consistency, confidence calibration, and structural validity of the extracted competency statements before the final augmented framework was used for module development (see Table 3). Coverage analysis showed that extracted evidence was concentrated in knowledge-related statements, reflecting the concept-heavy nature of job advertisements and implementation case descriptions. However, the retained evidence still represented all three KSA categories, indicating that the extraction process captured technical knowledge, operational skills, and higher-order abilities relevant to HRC training. Semantic consistency analysis indicated that KSA categories remained distinguishable, although some semantic overlap was observed because HRC competencies often combine technical, operational, cognitive, and managerial dimensions. Confidence calibration showed a negligible negative correlation between confidence ratings and evidence count ($r = -0.071$), indicating that confidence scores were not simply driven by frequency of occurrence. Structural validity analysis showed higher verb density in skill-related statements, supporting the expectation that skill evidence would contain more action-oriented language than knowledge or ability evidence.

External validation by three independent reviewers, separate from the Delphi panel, further supported the reliability of the augmented competency framework. The reviewers assessed classification accuracy, aggregation decisions, competency labeling, and KSA grouping. Agreement among reviewers was high, with an overall percent agreement of 93.3% and PABAK = 0.87, indicating strong agreement after adjustment for prevalence effects

Table 3: Internal validation of extracted HRC competency statements prior to final manual screening.

Validation dimension	Result
Statement coverage after duplicate removal	Knowledge = 419; Skills = 36; Abilities = 8; Total = 463
Final retained statements after manual review	Knowledge = 380; Skills = 30; Abilities = 8; Total = 418
Semantic consistency	KSA categories remained distinguishable, although some semantic overlaps were observed because HRC competencies often span technical, cognitive, and operational dimensions.
Confidence calibration	$r = -0.071$, indicating negligible correlation between confidence ratings and evidence count
Structural validity	Skills: $M = 1.14$; Knowledge: $M = 0.00$; Abilities: $M = 0.00$

4.2 Development of HRC training modules

4.2.1 Augmented competency input and construction robot-application contextual grounding

Consistent with the methodological design, the augmented HRC competency framework was used as the primary input for module development. The final input comprised 50 HRC competencies: 24 knowledge areas, 12 skills, and 14 abilities. Unique competency identifiers were retained throughout the pipeline to preserve traceability from competencies to cognitive classifications, LOs, module assignments, and Delphi validation results. Competencies were therefore treated as stable instructional constructs during module development. To provide contextual grounding, a complementary dataset mapping construction robot types to representative construction applications was developed (Table 4). The robot-application mapping ensured that competencies were situated within authentic construction robotics scenarios. By linking robot categories to task-specific site operations, the contextual mapping supported scenario-based LO generation and helped ensure that competency instantiation reflected realistic HRC conditions in construction practice. These robots and application contexts were later matched to competencies using semantic similarity, allowing the generated LOs to reference relevant construction robotics scenarios.

Table 4: Construction robot-application contextual grounding.

Construction Robots	Applications
Unmanned aerial vehicles (UAVs)	Aerial site monitoring; Surveying; Progress tracking; Site inspections; Aerial mapping
Bricklaying robots	Masonry and bricklaying
Gantry-based 3D printing robots	Additive construction and 3D printing
Industrial robotic arms	Sheathing; Tile laying; Drywall installation; Timber frame construction; Construction waste-sorting
Painting robots	Surface finishing and painting
Wearable robotic exoskeletons	Material handling/lifting support
Mobile terrestrial robots (e.g., Spot robot)	Ground-based site monitoring; Material transport
Autonomous construction layout robots	Layout mapping
Autonomous mobile robots	Safety monitoring; Reality capture for BIM/digital twins
Welding robots	Welding/joining
Construction cleaning robots	Construction site cleaning

4.2.2 Cognitive classification of HRC competencies

The hybrid rule-based Bloom mapping procedure classified the 50 HRC competencies across the KD and CPD of RBT. The classification used KSA-based default assignments and domain-specific keyword refinement rules, as described in the methodology. Each competency received one KD assignment and one CPD assignment, and an auditable assignment trace recorded whether the final cognitive level resulted from the KSA-based default assignment or keyword-based refinement. Across the KDs, HRC knowledge areas were classified as factual/conceptual knowledge, skills as procedural knowledge, and abilities as metacognitive knowledge. Across the CPD, competencies were distributed across Understand (15), Apply (9), Analyze (18), Evaluate (6), and Create (2). This distribution indicates that HRC training cannot be represented only as foundational robotics awareness; it also includes procedural implementation, analytical interpretation, evaluative judgment, and system-level design. Table 5 presents the joint distribution of competencies across KD and CPD. The complete list of HRC competencies and their cognitive classifications is provided in Supplemental Table S1.

Table 5: Cross-tabulation of HRC competencies across KD and CPD of RBT.

Knowledge dimension	Cognitive process dimension					Total
	Understand	Apply	Analyze	Evaluate	Create	
Factual/Conceptual	15	2	2	3	2	24
Procedural	0	7	3	2	0	12
Metacognitive	0	0	13	1	0	14
Total	15	9	18	6	2	50

4.2.3 Computationally supported generation of Bloom-aligned learning objectives

LOs were generated for each competency in the augmented HRC competency framework using a computationally supported process that combined the ABCD instructional design structure with Bloom-aligned action verbs. A total of 50 LOs were produced, each corresponding to one of the 50 HRC competencies. Each competency was mapped to a single LO, ensuring direct traceability between competency definitions and instructional outcomes. The cognitive levels of the generated LOs corresponded to the Bloom classifications assigned to the competencies. To enhance domain relevance, sentence embeddings were used to retrieve the top three semantically similar robot and application contexts for each competency from the construction robotics context dataset. These contexts were incorporated into the generated objectives, allowing each LO to be framed within realistic HRC situations. All LOs were generated using a Bloom-aligned verb canon, ensuring that the action verb used in each objective corresponds directly to the assigned cognitive process level (see Supplemental Table S2).

Internal validation checks confirmed that all objectives contained the appropriate Bloom-aligned verb and that no objectives were missing or incomplete. Examples of generated LOs include statements derived from specific competencies. For instance, for K4 – Sensors, the generated objective states: “Given a construction HRC scenario involving robotic sensing and inspection, the learner will be able to explain the functions, limitations, and safety implications of sensors used in collaborative robotic systems.” For K14 – System integration, the generated objective states: “Given a construction HRC scenario involving multiple robotic and digital systems, the learner will be able to design a system-integration strategy that addresses interoperability, safety, and performance requirements.” For S7 – Data analytics and management, the generated objective states: “Given robotic data from a construction monitoring scenario, the learner will be able to analyze data-management and analytics processes to support safety, productivity, or progress-related decisions.” For S11 – Robot operation and maintenance, the generated objective states: “Given a construction robotics operation scenario, the learner will be able to demonstrate appropriate robot operation and maintenance procedures consistent with task, safety, and performance requirements.” These examples illustrate how the generated objectives translated competencies into observable instructional outcomes with learner actions, construction HRC conditions, and performance criteria.

4.2.4 Theory-constrained HRC training module development

Following the generation of Bloom-aligned LOs, the competencies were organized into instructional modules through a theory-constrained module design procedure. In total, seven training modules were produced, covering the full set of 50 HRC competencies. Each competency was assigned to a single module, ensuring complete coverage of the competency framework. The resulting module structure reflects a progression from foundational knowledge of construction robotics to advanced competencies associated with system evaluation, safety management, and the deployment of robotic systems in construction environments. The modules begin with foundational knowledge of robotic systems and HRC principles in Module 1 (M1) and progressively address human–robot interfaces and interaction in Module 2 (M2), task planning and robot control in Module 3 (M3), modeling and simulation in Module 4 (M4), data analytics and robot learning in Module 5 (M5), safety and regulatory considerations in Module 6 (M6), and performance evaluation and deployment considerations in Module 7 (M7) (Table 6). This structure reflects a staged view of workforce preparation: learners first encounter foundational concepts and interaction mechanisms, then move toward procedural control, analytical modeling, data-driven reasoning, safety judgment, and deployment-oriented evaluation. The distribution of competencies across modules also reflects the integration of KSAs within each instructional unit. For instance, M1 primarily emphasized conceptual knowledge of robotic systems while incorporating robot operation and maintenance as a foundational operational skill. M2 integrated interaction-related knowledge, collaborative abilities, and interface-related skills. M3 combined task planning, control, programming, and system integration competencies. M4 brought together modeling, simulation, computational design, and systems thinking. M5 focused on data analytics,

machine learning, robot learning, and continuous learning. M6 integrated safety, ethics, regulation, standards, and decision-making. M7 addressed evaluation, ROI, project management, sustainability, adaptability, and robot acceptance and operational reliability.

Table 6: Theory-constrained organization of HRC competencies into training modules with cosine similarity scores.

Module	Competency category	Code	HRC competency	Number of competencies	Average cosine similarity
M1 - Foundations of Construction Robotics and Human-Robot Collaboration	Knowledge	K1	Types of robots	7	0.88
	Knowledge	K2	Construction robot applications		
	Knowledge	K20	ROS		
	Knowledge	K24	Robotics-related jobs		
	Knowledge	K3	Robot anatomy and technical specifications		
	Knowledge	K9	HRC-related fields		
	Skill	S11	Robot operation and maintenance		
M2 - Human-Robot Interfaces and Communication	Ability	A1	Teamwork	11	0.81
	Ability	A10	Spatial awareness		
	Ability	A11	Cultural and social awareness		
	Ability	A13	Robotic system reasoning		
	Ability	A2	Communication		
	Knowledge	K10	Immersive virtual environments		
	Knowledge	K11	Communication modes and technologies		
	Knowledge	K12	Human-robot interface		
	Knowledge	K4	Sensors		
	Skill	S1	Effective communication		
M3 - HRC Task Planning, Programming, and Control	Skill	S8	Human-robot interface proficiency	8	0.83
	Ability	A9	Critical thinking		
	Knowledge	K13	Robot control system		
	Knowledge	K14	System integration		
	Knowledge	K15	Programming		
	Knowledge	K5	Task planning		
	Skill	S2	Task planning proficiency		
	Skill	S5	Technical skill		
M4 - Modeling, Simulation, and Systems Thinking	Skill	S6	Programming proficiency	7	0.85
	Ability	A4	Problem-solving		
	Ability	A6	Attention to detail		
	Ability	A7	Analytical aptitude		
	Knowledge	K16	Modeling and simulation		
	Knowledge	K19	Computational design		
	Skill	S10	Simulation and modeling		
M5 - Data Analytics, Machine Learning, and Robot Learning	Skill	S12	Systems thinking	5	0.88
	Ability	A3	Continuous learning		
	Knowledge	K17	Data analytics and machine learning		
	Knowledge	K18	Robot learning methods		
	Skill	S7	Data analytics and management		
M6 - Safety, Ethics, Regulation, and Standards	Skill	S9	Application of machine learning algorithms	6	0.87
	Ability	A12	Safety awareness		
	Ability	A8	Decision-making		
	Knowledge	K6	HRC ethics and regulation		
	Knowledge	K7	HRC safety and standards		

Module	Competency category	Code	HRC competency	Number of competencies	Average cosine similarity
M7 - Evaluation, Performance, ROI, and Deployment	Skill	S3	Regulation standards compliance	6	0.81
	Skill	S4	Safety management		
	Ability	A14	Robot acceptance and operational reliability		
	Ability	A5	Adaptability		
	Knowledge	K21	Cost and ROI in robotics		
	Knowledge	K22	Robotic project management		
	Knowledge	K23	Robotics sustainability		
	Knowledge	K8	HRC evaluation		

4.3 Evaluation of developed HRC training modules

4.3.1 Computational evaluation of HRC training modules

Before human evaluation, the developed training modules underwent an internal computational evaluation to assess structural consistency and completeness. All 50 competencies were successfully assigned valid Bloom cognitive levels, translated into ABCD-informed learning objectives, and allocated to one of the seven instructional modules without missing assignments. Computational checks further confirmed that within-module semantic coherence analysis demonstrated strong internal consistency across all modules, with mean intra-module similarity scores ranging from 0.81 to 0.88. This result supports the coherence of the theory-constrained assignment strategy: although modules were not created by unsupervised clustering, their final competency compositions were semantically coherent. The modules contained between 5 and 11 competencies, satisfying the instructional design constraint of maintaining manageable module sizes while preserving balanced coverage of the overall competency framework. To further examine instructional coverage and cognitive distribution across the training architecture, a competency–module coverage matrix was developed (Supplemental Tables S3A– S3C). The matrix mapped each HRC competency to its assigned instructional module and represented the Bloom-based cognitive depth associated with the competency within the module structure. The coverage matrix confirmed that all 50 competencies were represented across the seven-module architecture and demonstrated a balanced distribution of foundational, procedural, analytical, evaluative, and higher-order competencies across the instructional sequence. The matrix also provided an auditable linkage between competency augmentation, cognitive classification, LO generation, and module assignment.

4.3.2 Delphi validation of HRC training modules

The Delphi study validated the developed HRC training modules across five dimensions, including competency relevance, module–competency fit, cognitive classification, LO quality, and module completeness. These dimensions were evaluated in Round 1, and non-consensus or ambiguous items were refined and reassessed in Round 2 using aggregated quantitative ratings and qualitative feedback from the expert panel.

Relevance of competency to the assigned module

The relevance of HRC competencies was evaluated across all seven training modules in Round 1, with non-consensus items re-evaluated in Round 2. Strong agreement was observed for core technical and operational competencies, with consensus items typically ranging from 73% to 100% agreement (Table 7). M3–M7 achieved consensus for all assigned competencies, confirming that competencies in task planning, programming, robot control, modeling and simulation, systems thinking, data analytics, robot learning, safety and regulation, evaluation and deployment were consistently perceived as relevant to HRC training. M1 and M2 exhibited localized disagreement concentrated in a small subset of competencies. In M1, K20, K24, and S11 did not achieve consensus after re-evaluation. In M2, A11, A13, K10, and K4 remained below consensus despite moderate improvements between rounds. These patterns indicate that disagreement was concentrated in boundary-spanning, emerging, or socio-cognitive competencies rather than distributed across the full module framework. Detailed median ratings, percentage ratings of 4 or higher, and consensus status across the first and second rounds are provided in Supplemental Tables S4 and S5.

Table 7: Relevance of HRC competency across modules.

Module	Total competencies	Consensus achieved in round 1	Consensus achieved in round 2	Final consensus (%)	Non-consensus competencies
M1	7	4	4	57	K20, K24, S11
M2	11	6	7	64	A11, A13, K10, K4
M3	8	8	8	100	None
M4	7	7	7	100	None
M5	5	4	5	100	None (A3 resolved)
M6	6	6	6	100	None
M7	6	6	6	100	None

Conceptual appropriateness between HRC modules and competencies

The conceptual appropriateness of the module–competency structure was evaluated by asking experts whether each HRC competency was appropriately assigned to its module. High agreement was observed across most modules, with consensus levels typically exceeding 80% (Table 8). M3–M7 demonstrated complete consensus, indicating clear conceptual boundaries for modules related to task planning and control, modeling and systems thinking, data analytics, safety and regulation, and evaluation and deployment. Localized ambiguities were observed in M1 and M2. In M1, K20 and K3 achieved consensus after Round 2 reassessment, while S11 remained below the threshold, indicating persistent ambiguity in the placement of robot operation and maintenance. In M2, initial concerns around A11, K10, and K4 were resolved in Round 2, resulting in full consensus. These results indicate that M1 retained one unresolved boundary-spanning competency, while initial module-boundary concerns in M2 were resolved after Round 2 reassessment. The results suggest that module–competency misalignment was primarily due to boundary-spanning HRC competencies rather than structural weakness in the seven-module architecture. Details of the percentage of yes and no responses, the dominant suggested module, and consensus status across the first and second rounds are provided in Supplemental Tables S6 and S7.

Table 8: Module-competency conceptual fit across modules.

Module	Total	Consensus achieved in round 1	Consensus achieved in round 2	Final consensus (%)	Key reassignment suggestions
M1	7	4	6	86	K20 → M3; S11 → M2/M5/M3
M2	11	8	11	100	A11 → M6; K10 → M4; K4 → M5
M3	8	8	8	100	None
M4	7	7	7	100	None
M5	5	5	5	100	None
M6	6	6	6	100	None
M7	6	6	6	100	None

Assessment of cognitive classification of HRC competencies

The cognitive classification of HRC competencies was evaluated by assessing the appropriateness of assigning each competency to the KD and CPD of RBT. Overall, strong alignment was observed across modules, with most competencies achieving consensus in Round 1 and all remaining discrepancies resolved in Round 2 (Table 9). Across modules, agreement ranged from 73% to 100%, indicating convergence in expert interpretation of cognitive demand.

Initial disagreement was concentrated in competencies involving procedural execution, system integration, programming, safety management, systems thinking, ROI, and deployment evaluation, where boundaries between procedural, analytical, and evaluative reasoning are inherently fluid. In M1, K20 was the only item that did not achieve consensus initially, with experts suggesting classification under procedural knowledge at the “Apply” level. After revision, the percentage agreement increased in Round 2, achieving consensus. In M2, A11 showed borderline agreement due to uncertainty between factual/conceptual knowledge and the “Understand” level, but consensus was achieved in the second round. Disagreement was observed in M3, where K14 and K15 did not initially reach consensus in the first round because experts differed between procedural knowledge at the “Apply” and “Create” levels. However, agreement was achieved after revision in the second. In M4, S12 initially showed low agreement because experts associated it with metacognitive knowledge at the “Evaluate” level; it reached full

agreement in the second round. Similarly, S4 in M6 achieved consensus after being clarified as procedural knowledge at the “Apply” level. The largest disagreement occurred in M7, where only 2 of 6 competencies achieved consensus in the first round. Items such as A14, A5, K21, and K8 showed ambiguity between procedural and metacognitive knowledge and between the “Analyze” and “Evaluate” levels. After refinement, all competencies achieved consensus in the second round. Details of the cognitive classification results are provided in Supplemental Tables S8 and S9.

Table 9: Cognitive classification of HRC competencies across modules.

Module	Total competencies	Consensus in round 1	Consensus in round 2	Final consensus (%)	Key cognitive issues
M1	7	6	7	100%	K20 (Procedural–Apply ambiguity)
M2	11	10	11	100%	A11 (Understand vs higher-level ambiguity)
M3	8	6	8	100%	K14, K15 (Apply vs Create)
M4	7	6	7	100%	S12 (Evaluate/metacognitive)
M5	5	5	5	100%	None
M6	6	5	6	100%	S4 (Apply-level clarification)
M7	6	2	6	100%	Evaluate/Analyze ambiguity (A14, K21, K8)

Table 10: Summary interpretation of LO quality across modules based on Delphi ratings.

Module	Clarity	Measurability	Usability	Persistent issues
M1	High	Low	Moderate → High	K20, S11 measurability
M2	Moderate	Low	High	S1, A1, A2, A11 measurability
M3	Moderate	Moderate	High	A9 clarity/measurability
M4	Moderate → High	Low	High	S12 measurability
M5	Moderate	Low	High	A3 unresolved
M6	Moderate	Moderate	High	A12, A8 measurability
M7	High	Moderate → High	High	Initially A14, A5

Table 11: Cross-tabulation of learning-objective measurability consensus by KSA type across Delphi rounds.

KSAtType	Total competencies	Round 1 consensus, n (%)	Round 2 consensus, n (%)	Persistent non-consensus, n (%)
Knowledge	24	14 (58.3%)	23 (95.8%)	1 (4.2%)
Skills	12	8 (66.7%)	9 (75.0%)	3 (25.0%)
Abilities	14	0 (0.0%)	7 (50.0%)	7 (50.0%)
Total	50	22 (44.0%)	39 (78.0%)	11 (22.0%)

Evaluation of learning objective quality

LO quality was evaluated across clarity, measurability, and usability. Overall, usability showed consistently high agreement across modules, while clarity achieved moderate-to-high agreement. In contrast, measurability remained the most challenging dimension, with several HRC competencies failing to reach consensus even after refinement (Table 10). Consensus was generally achieved for clarity and usability (73%–100%), indicating that most LOs were understandable and instructionally applicable. However, measurability frequently remained below the 70% threshold, particularly for higher-order, socio-cognitive, adaptive, and judgment-based competencies. The cross-tabulation by KSA type (Table 11) indicates that this was a systematic characteristic of ability-oriented competencies rather than isolated shortcomings of individual learning objectives. While most knowledge and skill competencies were successfully translated into measurable learning objectives following iterative refinement, ability competencies remained substantially more difficult to operationalize, with persistent non-consensus affecting 50.0% of ability competencies compared with 25.0% of skills and 4.2% of knowledge competencies. This pattern suggests that the principal challenge in HRC instructional design lies not in defining technical knowledge or procedural skills, but in translating abilities involving socio-cognitive, adaptive, relational, or judgment-based constructs (such as communication, decision-making, safety awareness, teamwork, continuous learning, critical thinking, and cultural and social awareness) into observable and assessable instructional outcomes. M2–M4 showed strong usability and moderate clarity, but measurability remained inconsistent for

competencies such as A1, A2, A9, S1, and S12, reflecting difficulties in defining observable performance criteria for interaction-based and systems-oriented reasoning tasks. In M5, most competencies achieved consensus; however, although A3 achieved consensus regarding its relevance to the module (Table 7), the associated LO remained unresolved in Table 10 due to low measurability ratings, suggesting that adaptive and lifelong learning constructs are difficult to translate into assessable objectives. Similarly, in M6, A8 and A12 continued to show low clarity and measurability due to their context-dependent and judgment-based nature. In contrast, M7 demonstrated strong performance across all dimensions after refinement, with measurability for A14 and A5 improving to consensus levels. Supplemental Tables S10 and S11 provide detailed ratings, agreement percentages, and consensus outcomes across both Delphi rounds. Rather than indicating failure of the module framework, these persistent measurability issues identify the types of HRC competencies that require more flexible and authentic assessment strategies, such as performance rubrics, scenario-based evaluation, reflective judgment tasks, simulation-based demonstrations, deployment plans, dashboards, compliance audits, checklists, and lab-based artifacts.

Completeness of developed HRC training modules

Module completeness was evaluated to assess the extent to which the developed training modules collectively provided adequate coverage of HRC competencies and maintained structural coherence across KSAs. The results indicate strong consensus across all evaluation items, with median ratings of 4 or higher and agreement levels ranging from 80% to 100% (Table 12). The overall coverage of competencies required for effective learning and practice of HRC in construction achieved the highest level of agreement, with a median rating of 5 and 100% agreement. Strong agreement was also observed for coverage of key HRC KSAs, balance across KSA domains, alignment between competencies and LOs, and module sequencing. These findings indicate that the proposed modules were collectively complete in terms of domain coverage, KSA balance, LO alignment, and instructional sequencing. This module-level consensus coexisted with localized concerns about selected competencies and LO measurability, suggesting that the framework is broadly coherent while still requiring refinement for specific complex competencies.

Table 112: Module completeness evaluation.

Evaluation item	Median rating	Percentage rating ≥ 4	Consensus status
Coverage of key HRC KSAs across modules	4	87%	Achieved
Coverage of competencies needed for effective HRC learning and practice	5	100%	Achieved
Balance between KSAs	4	87%	Achieved
Alignment of competencies and LOs with modules	4	87%	Achieved
Sequence and ordering of modules	4	80%	Achieved

Qualitative feedback from Delphi participants

Qualitative expert feedback complemented the quantitative Delphi findings by explaining the patterns of agreement and disagreement. Experts generally supported the alignment between HRC competencies and instructional modules while emphasizing curriculum sequencing, foundational robotics knowledge, construction contextualization, and progression from conceptual understanding to advanced technical implementation.

Several experts distinguished between introductory conceptual competencies and implementation-oriented skills. Competencies related to ROS, robot operation and maintenance, system integration, programming, and advanced technical implementation were viewed by some experts as more appropriate for later instructional stages or specialized technical modules. This feedback helps explain the persistent relevance and module-fit disagreements surrounding K20 and S11. In particular, S11 was perceived by some experts as foundational because learners should understand robot operation and maintenance early, while others viewed it as an applied operational skill better positioned after learners acquire interface, task-planning, data, or safety competencies. Experts also emphasized the importance of construction-contextualized learning. They noted that HRC training should connect robotic systems to practical construction applications, site constraints, worker safety, project controls, productivity, and deployment decisions. This theme supports the study's use of robot–application contextual grounding during LO generation and reinforces the need for construction-specific examples in future instructional implementation. A third theme concerned cross-cutting competencies. Experts identified safety awareness, critical thinking,

analytical reasoning, communication, decision-making, and systems thinking as competencies that extend across multiple modules rather than belonging exclusively to a single instructional unit. This finding explains why some competencies produced module-boundary ambiguity despite the overall coherence of the seven-module structure. It also suggests that future course implementation may need to treat certain competencies as recurring threads across modules rather than as isolated, module-specific outcomes. Qualitative feedback also supported the cognitive classification of HRC competencies using RBT while identifying areas requiring refinement. Experts emphasized the distinction between conceptual understanding and procedural execution. Several participants suggested that competencies related to ROS, programming, system integration, robot operation, safety management, and machine learning implementation should be mapped to procedural knowledge and Apply- or Analyze-level cognitive processes rather than lower-order understanding alone. Participants also highlighted the importance of higher-order cognition in HRC learning, particularly for competencies involving troubleshooting, systems reasoning, decision-making, ethics, safety, deployment, and ROI evaluation.

Feedback on LOs reinforced the quantitative finding that measurability was the weakest dimension. Experts recommended clearer performance conditions, measurable criteria, assessment rubrics, performance thresholds, and tangible deliverables such as reports, deployment plans, dashboards, compliance audits, checklists, inspection records, safety reviews, and lab-based artifacts. Detailed thematic analysis, including dimensions, themes, and codes, is provided in Supplemental Table S12.

5. DISCUSSION

5.1 From HRC competency identification to training design

This study addressed a critical gap in construction robotics workforce development by moving beyond competency identification toward a process for translating HRC competencies into validated training modules. Prior research has shown that construction robotics requires new workforce competencies related to robot operation, sensing, task coordination, human–robot interfaces, safety, communication, and decision-making (Bademosi & Issa, 2022; Olukanni et al., 2025). However, identifying competencies alone does not explain how those competencies should be organized, taught, assessed, or aligned with the management responsibilities of civil engineering and construction organizations. This study contributes to this gap by demonstrating a traceable pathway from competency augmentation to cognitive classification, LO development, module formation, and Delphi-based expert validation. The resulting framework is intended to support learners and professionals who are expected to deploy, supervise, coordinate, or collaborate with robotic systems in construction rather than robotics specialists responsible for developing robotic hardware or control algorithms. Accordingly, the framework is applicable to senior undergraduate and graduate students in construction-related disciplines, as well as early-career construction engineers, field engineers, superintendents, project managers, safety professionals, BIM/VDC personnel, and other practitioners preparing to work in human–robot collaborative environments. The findings also show that HRC workforce training should be understood as a competency-to-training translation problem. The study began with a 42-competency framework and expanded it to 50 competencies using industry-facing evidence from job advertisements and robotic implementation case studies. The augmented competencies were then cognitively classified, translated into ABCD-informed LOs, organized into seven theory-constrained modules, and evaluated by experts across relevance, module–competency fit, cognitive classification, LO quality, and module completeness. Because the framework targets a broad construction workforce rather than a single occupational group, it is sufficiently flexible to support multiple educational and professional contexts. In higher education, the modules can serve as the basis for undergraduate and graduate courses on construction robotics and HRC, while in industry they can support onboarding, continuing professional development, and workforce upskilling for personnel involved in planning, construction, inspection, commissioning, and facility operations. Although the competency framework remains consistent across these contexts, the instructional depth, assessment strategies, and delivery methods should be adapted to the prior knowledge, responsibilities, and experience of the intended learner population. This progression is important because HRC training is not only a technical education problem but also a workforce planning, technology management, safety governance, and organizational readiness problem.

The overall findings demonstrate the coherence and validity of the framework. Experts reached a strong consensus on module completeness, cognitive classification after refinement, and module–competency fit for most modules. However, the results also show that HRC training cannot be reduced to rigid module categories. Some competencies remained difficult to place or measure because they cross technical, social, operational, and

managerial boundaries. This contrast is central to the study's contribution: the framework is sufficiently structured to support curriculum and workforce planning, but flexible enough to reveal where HRC competencies require cross-module reinforcement, authentic assessment, and expert judgment. The interpretation of these findings should also consider the composition of the Delphi panel. Although the predominance of academic experts was appropriate given the study's emphasis on competency structure, cognitive classification, learning objective development, and instructional coherence, it may have placed greater emphasis on pedagogical organization than on the operational realities of workforce training in construction. Practical constraints such as project schedules, production demands, equipment availability, supervisory practices, organizational training resources, and site-specific implementation conditions were represented through practitioner participation but were not the dominant perspective. Consequently, the proposed framework should be regarded as an expert-validated instructional architecture that would benefit from additional practitioner-led validation involving contractors, craft workers, safety professionals, equipment operators, and organizational training leaders before widespread industry implementation. Furthermore, the framework demonstrated strong expert consensus regarding competency organization, cognitive progression, and instructional design, its practical educational effectiveness has not yet been empirically evaluated with learners. Consequently, the findings should be interpreted as evidence of a validated instructional design framework rather than evidence of improved learning outcomes or workforce performance. Empirical studies involving students and construction practitioners are needed to evaluate learning gains, competency development, transfer to practice, and usability in authentic HRC training environments. The following sections discuss what the augmented competencies reveal about workforce readiness, how the modules organize those competencies, where assessment challenges remain, and how the framework can support educational programs and construction organizations.

5.2 Expanded competency demands for robotics adoption and workforce readiness

The competency augmentation results show that HRC workforce preparation must extend beyond traditional technical robotics content. The eight newly identified competencies—cost and ROI in robotics, robotic project management, robotics sustainability, robotics-related jobs, robot operation and maintenance, systems thinking, robotic system reasoning, and robot acceptance and operational reliability—broaden the framework from technical capability toward economic, organizational, sustainability, operational, and adoption-oriented readiness. These additions are important because robotic systems are adopted within project-based organizations where cost justification, safety, reliability, workforce roles, and implementation planning influence whether technologies are used successfully. This finding supports prior construction automation and robotics adoption research. T. Bock (2015) argued that construction automation is moving toward broader robotic integration, while Davila Delgado et al. (2019) showed that adoption barriers in construction include contractor-side economic factors, client-side economic factors, technical and work-culture factors, and weak business-case factors. The present findings extend this literature by identifying workforce competencies that can help organizations respond to those adoption barriers. For example, cost and ROI in robotics directly address the need to evaluate economic value; robotic project management supports planning and coordination; robot acceptance and operational reliability addresses trust, dependability, and sustained use; and systems thinking supports integration across tasks, technologies, and project constraints.

The findings also qualify the role of computational competency augmentation. Industry-facing sources are valuable because they reveal emerging terminology, current workforce expectations, and deployment-oriented concerns that may not yet be fully represented in academic literature. However, job advertisements and implementation case studies may overrepresent visible technologies, vendor language, and current market needs while underrepresenting deeper socio-cognitive competencies. This pattern was reflected in the extracted evidence, where knowledge-related statements were much more frequent than skill- and ability-related statements. Therefore, computational augmentation should be treated as a complementary evidence source, not as a substitute for theory, expert judgment, or instructional design. This supports the human-in-the-loop structure used in the study, where extracted and LLM-mapped candidate competencies were manually screened and independently reviewed before being incorporated into the final framework.

5.3 Module coherence and boundary-spanning competencies

The Delphi results provide evidence that the seven-module structure is conceptually coherent and instructionally meaningful. M3–M7 achieved strong consensus across relevance and module–competency fit, indicating that

experts viewed task planning, programming and robot control, modeling, simulation and systems thinking, data analytics and robot learning, safety and regulation, and HRC evaluation and robot deployment as stable instructional domains. The module completeness results also showed strong agreement for KSA coverage, module sequencing, and alignment between competencies and LOs. These findings support the structural coherence of the proposed HRC training architecture. However, the results also show that some HRC competencies are inherently boundary-spanning, consistent with prior research on boundary-spanning roles that operate across multiple functional domains (Levina & Vaast, 2005). M1 and M2 showed the most disagreement, especially for ROS, robotics-related jobs, robot operation and maintenance, immersive virtual environments, sensors, cultural and social awareness, and robotic system reasoning. These competencies are not necessarily misplaced; rather, they can reasonably belong to more than one instructional context. For example, sensors can be introduced as part of foundational robotics, interface design, data analytics, or safety monitoring. Robot operation and maintenance can be treated as foundational exposure, an applied technical skill, a safety competency, or a deployment-readiness competency. Cultural and social awareness may be positioned within communication, safety, team coordination, or organizational adoption. This finding both supports and complicates modular curriculum design. It supports modularization because most modules achieved high consensus and semantic coherence. It complicates modularization because HRC practice does not always follow clean disciplinary boundaries. For civil engineering and construction organizations, this means HRC modules should be treated as primary instructional homes rather than exclusive containers. Boundary-spanning competencies may need to appear as recurring threads across modules, assignments, and assessments. For example, robot operation and maintenance could be introduced in foundational robotics, practiced in interface and control modules, and assessed in safety or deployment scenarios. This approach would preserve the clarity of the seven-module structure while acknowledging the integrated nature of construction robotics practice.

5.4 Cognitive progression and assessment design

The cognitive classification results show that HRC competencies span conceptual, procedural, analytical, evaluative, and design-oriented demands. This is important because HRC training cannot be limited to awareness of robotic technologies. Learners must understand robotic systems, apply procedures, analyze data and system behavior, evaluate safety and deployment trade-offs, and design or adapt solutions for construction contexts. The broad distribution across Understand, Apply, Analyze, Evaluate, and Create levels confirms that HRC training requires progressive cognitive development rather than single-level technical exposure.

The results support the usefulness of RBT as a design scaffold for HRC curriculum development. RBT distinguishes KD from CPD, allowing instructional designers to differentiate factual/conceptual knowledge, procedural performance, and metacognitive reasoning (Anderson & Krathwohl, 2001; Krathwohl, 2002). In this study, the RBT-based procedure helped translate KSA categories into cognitive expectations and then into LOs and modules. The full resolution of cognitive-classification discrepancies in Round 2 suggests that the approach was acceptable to experts after refinement. At the same time, the results caution against the mechanical use of Bloom-based classification. Initial disagreements were concentrated in competencies such as system integration, programming, systems thinking, safety management, ROI evaluation, and deployment-oriented decision-making. These competencies often require more than one cognitive process. For example, system integration may require understanding system components, applying integration procedures, analyzing interoperability issues, evaluating safety and performance constraints, and designing improved configurations. Thus, RBT should be used as an adaptive instructional design tool rather than as a rigid coding scheme. In complex sociotechnical domains such as HRC, expert review is needed to ensure that cognitive classifications reflect how competencies are actually applied in construction settings. The LO results extend this interpretation by showing that measurability is the most important instructional-design challenge in HRC training. Experts generally found the LOs usable, and clarity was moderate to high; however, the cross-tabulation by KSA type demonstrated that measurability challenges were not randomly distributed across competencies. Rather, they were concentrated among ability competencies, with persistent non-consensus affecting 50.0% of ability competencies after iterative refinement. This finding indicates that the primary challenge in competency-based HRC curriculum development is not translating technical knowledge or procedural skills into learning objectives, but transforming higher-order socio-cognitive capabilities into observable and assessable outcomes. Competencies such as teamwork, communication, cultural and social awareness, continuous learning, decision-making, safety awareness, and systems thinking are fundamental to effective HRC practice, yet they depend on coordination, professional judgment, adaptation, and contextual

reasoning that are difficult to capture through conventional objective-based assessments. Consequently, these competencies require assessment approaches that extend beyond knowledge tests or procedural checklists. Authentic assessment strategies, including simulation-based exercises, scenario-based problem solving, structured observation, performance rubrics, reflective journals, team-based activities, workplace demonstrations, and instructor or supervisor evaluations, are better suited to capturing the adaptive and socio-cognitive capabilities required for effective HRC. Instead of applying a uniform assessment approach across all competency types, competency-based HRC training should align assessment methods with the nature of the competency being evaluated. Knowledge-oriented competencies may be assessed through conventional examinations or quizzes, procedural skills through practical demonstrations and performance tasks, whereas higher-order ability competencies require authentic, context-rich assessments that evaluate judgment, collaboration, adaptability, and decision-making in realistic construction environments. This finding aligns with CBE, which emphasizes demonstrable outcomes rather than time-based exposure or content coverage (Perera et al., 2017). It also aligns with objective-writing principles that require observable behaviors, conditions, and performance criteria (Mager & Peatt, 1962).

The initial KSA-to-RBT mapping should also be interpreted as a computational starting heuristic rather than a definitive cognitive classification. Consistent with CBE and the RBT, the heuristic provided an initial association between knowledge, skills, and abilities and the corresponding knowledge dimensions, while recognizing that many HRC competencies integrate declarative understanding, procedural execution, analytical reasoning, and self-regulation (Anderson & Krathwohl, 2001; Krathwohl, 2002; Mulder, 2014). This was particularly evident for competencies in M7, where expert disagreement centered on the distinction between procedural and metacognitive knowledge and between analytical and evaluative reasoning. Contrary to the indication of a limitation in the framework, these disagreements illustrate the inherently integrative nature of advanced HRC competencies and reinforce the importance of combining computational heuristics with domain-specific refinement rules and iterative expert validation.

In practice, these higher-order competencies should be assessed using authentic performance-based approaches that reflect the collaborative and dynamic nature of HRC in construction. For example, trust can be evaluated through repeated human–robot collaboration scenarios that examine appropriate reliance on robotic systems under varying levels of autonomy and operational uncertainty. Communication competencies are more appropriately assessed through structured observation rubrics during collaborative tasks involving workers, supervisors, and robotic systems, where information exchange and coordination can be directly observed. Safety judgment can be assessed through scenario-based hazard identification, risk assessment, and incident response exercises that require learners to justify safety-related decisions in realistic construction contexts. Similarly, systems thinking is best evaluated through integrated project cases that require learners to analyze interactions among robotic systems, construction workflows, safety requirements, and project objectives, using deliverables such as deployment plans, system models, concept maps, or reflective analyses. These assessment approaches provide observable evidence of complex competency attainment while remaining consistent with competency-based education and the sociotechnical nature of HRC.

5.5 Computationally supported and human-in-the-loop curriculum development

The study demonstrates how computational methods can support curriculum development without replacing domain expertise. The workflow used PDF extraction, OCR fallback, HRC-context filtering, rule-based KSA classification, Sentence-BERT semantic clustering, local LLM-assisted candidate mapping, RBT-based cognitive profiling, contextual grounding, module assignment, module-coherence checking, and coverage assessment. These tools improved traceability and scalability, but their value depended on manual screening, independent reviewer validation, and Delphi-based expert judgment.

This human-in-the-loop design is important for two reasons. First, computational methods can efficiently process large amounts of unstructured evidence, but they can also reproduce source biases or overemphasize frequent terminology. Sentence-BERT supports semantic comparison by producing sentence embeddings that can be compared using cosine similarity (Reimers & Gurevych, 2019), but semantic similarity alone does not determine instructional relevance. Second, LLM-assisted mapping can help compare extracted evidence against a baseline framework, but it may also overgenerate categories or produce plausible but overly broad labels. Therefore, the LLM in this study was appropriately used as a candidate-mapping aid rather than as an autonomous curriculum

generator or validator. For construction engineering and management research, this is a key methodological implication. Research and practice increasingly use AI, NLP, and analytics to support decision-making, but workforce and curriculum decisions require accountability, interpretability, and stakeholder relevance. The study contributes a responsible computational workflow in which AI-supported extraction and mapping are embedded within instructional theory and expert validation. This approach can be adapted to other emerging construction technology domains where competency needs evolve quickly, but where training modules must remain interpretable, assessable, and aligned with professional practice.

5.6 Practical implications for civil engineering and construction management

The findings have direct practical implications for academic programs, industry training providers, and construction firms adopting robotic systems. First, the seven-module framework provides a curriculum-mapping tool for civil engineering, construction engineering, and construction management programs. Programs can use the modules to identify where HRC content already appears in existing courses and where new content is needed. For example, foundational robotics and HRC concepts may fit introductory construction technology courses, while data analytics, safety, simulation, and deployment evaluation may fit advanced construction automation, digital construction, safety, or project management courses. Second, the framework can support continuing education and workforce upskilling. Construction organizations adopting drones, layout robots, mobile robots, exoskeletons, robotic arms, or automated inspection systems need more than equipment-specific training. They need workers and managers who can interpret robotic outputs, coordinate work around robotic systems, evaluate safety and reliability, assess cost and value, and understand when robotics should or should not be used. The augmented competencies related to ROI, robotic project management, sustainability, operational reliability, and robot acceptance are particularly relevant because they connect robotics training to implementation decisions. Third, the framework can guide role definition and capability-gap analysis. Field personnel may require competencies in sensing, robot operation, safety management, interface use, and task execution. Project managers and technology champions may require competencies in ROI, deployment planning, system integration, sustainability, and operational reliability. Educators and trainers can use these distinctions to design role-specific learning pathways rather than one-size-fits-all robotics training. Fourth, the findings point to the need for performance-based assessment in HRC training. Because measurability was the weakest LO dimension, educators and trainers should avoid relying only on written quizzes or generic objective statements. More appropriate assessment formats may include robotic operation labs, task-planning scenarios, simulation-based troubleshooting, dashboard interpretation, safety-risk assessments, deployment proposals, compliance audits, and ROI evaluations. These assessment formats would connect competency attainment to observable construction robotics performance. Finally, the framework can help reduce workforce–technology misalignment. Robotics adoption can slow when organizations deploy technologies faster than they prepare workers to use, supervise, evaluate, and adapt those technologies. By linking competencies to LOs, modules, and construction robotics contexts, the framework enables training to be planned alongside technology deployment. This implication is consistent with construction robotics adoption research showing that weak business cases, economic uncertainty, technical challenges, and work-culture barriers can limit implementation (Davila Delgado et al., 2019). In this sense, the framework supports not only education but also technology-adoption planning, workforce readiness, and organizational learning.

6. CONCLUSION

This study developed and expert-validated a data-informed and theory-guided framework for translating HRC competencies into training modules for the construction workforce. By integrating PDF-based source extraction, rule-based KSA classification, Sentence-BERT semantic clustering, local LLM-assisted candidate mapping, RBT-based cognitive profiling, ABCD-informed LO generation, embedding-supported contextual grounding, theory-constrained module formation, module-coherence checking, coverage assessment, and Delphi-based expert validation, the study provides a transparent and reproducible workflow for HRC workforce training design. The framework expanded a baseline set of 42 HRC competencies to 50 competencies and organized them into seven modules that support progression from foundational robotics knowledge and human–robot interfaces to task planning, modeling and simulation, data analytics, safety and regulation, and deployment-oriented evaluation. The findings show strong expert support for module–competency fit, cognitive classification, module completeness, and overall instructional coherence. At the same time, selected boundary-spanning, socio-cognitive, adaptive, and higher-order competencies remained difficult to operationalize as measurable LOs. This finding demonstrates that

the key challenge in HRC workforce training is not only identifying relevant competencies, but also translating complex human–robot collaboration capabilities into observable, teachable, and assessable learning outcomes. The framework contributes to construction engineering and management by offering an approach for aligning construction robotics adoption with workforce planning, curriculum development, training standardization, assessment design, and organizational readiness. The study also shows that HRC training should be treated as part of technology and innovation management in construction organizations. The inclusion of competencies related to cost and ROI, robotic project management, sustainability, robot acceptance, operational reliability, systems thinking, and robot operation and maintenance indicates that the adoption of robotics requires more than technical proficiency. It also requires workers, managers, educators, and training providers who can evaluate value, manage risks, coordinate human–robot work, support safety governance, and plan for long-term technology integration.

This study has several limitations that should guide future research. First, the study focused on module development and expert validation and did not evaluate instructional effectiveness in classroom, industry, or field-training settings. Therefore, the findings do not yet establish learning gains, transfer to practice, or changes in workforce performance. Second, competency augmentation depended on selected literature-derived competencies and industry-facing documentary sources, which may influence the scope of the resulting framework. Third, although the computational workflow supported source extraction, KSA classification, semantic clustering, and local LLM-assisted candidate mapping, final competency interpretation depended on manual screening and expert judgment. Fourth, although the Delphi panel intentionally combined academic and industry expertise, the predominance of academic experts may limit the extent to which the validated framework reflects the operational realities of construction workforce training. Broader practitioner-led validation involving contractors, craft workers, safety professionals, equipment operators, and organizational training leaders is therefore warranted.

Future research should implement and evaluate the expert-validated modules in construction education and industry training settings to assess learning effectiveness, usability, learner engagement, transfer to practice, and workforce readiness for robotics adoption. Future work should also develop performance rubrics, simulation-based assessments, scenario-based tasks, and artifact-based evaluations to improve the measurability of socio-cognitive and higher-order HRC competencies. In addition, the modules could be extended into digital, immersive, and adaptive learning environments, including virtual reality, augmented reality, and web-based platforms. Future studies should also examine how module relationships and competency progression can support adaptive sequencing, personalized learning pathways, and continuous updating of HRC training as robotic technologies and construction practices evolve. Finally, future research should investigate how module sequencing and instructional strategies can be tailored to different learner groups, recognizing that the intended learning level, instructional depth, assumed prior knowledge, delivery methods, and assessment approaches vary substantially among undergraduate students, graduate students, field workers, project managers, safety personnel, and BIM/VDC or robotics professionals.

DISCLOSURE STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation (NSF) under Grant Nos. 2235375 and 2402008.

SUPPLEMENTAL MATERIALS

Supplemental tables, codes, input data, and generated outputs associated with the HRC competency augmentation, theory-constrained HRC training module development framework, and Delphi validation underlying this manuscript are made available through the [Virginia Tech Data Repository](https://doi.org/10.7294/32394888) at <https://doi.org/10.7294/32394888>.

REFERENCES

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc.
- Bademosi, F. M., & Issa, R. R. (2022). Essential knowledge, skills, and abilities required for talent cultivation in



- construction automation and robotics. In *Automation and robotics in the architecture, engineering, and construction industry* (pp. 31–57). Springer. https://doi.org/10.1007/978-3-030-77163-8_2
- Barasa, L., Nopriani, Y., & Yoniessa, S. (2026). NLP-Powered Curriculum Material Generation for Multilingual Maritime Technical Education. *IJISIT: International Journal of Computer Science and Information Technology*, 3(1), 76–83. <https://doi.org/10.55123/ijisit.v3i1.65>
- Bock, T. (2015). The future of construction automation: Technological disruption and the upcoming ubiquity of robotics. *Automation in Construction*, 59, 113–121. <https://doi.org/10.1016/j.autcon.2015.07.022>
- Bock, T., & Linner, T. . (2015). Advanced Construction and Building Technology. In T. Bock & T. Linner (Eds.), *Robot-Oriented Design: Design and Management Tools for the Deployment of Automation and Robotics in Construction* (pp. 1–17). Cambridge University Press. <https://doi.org/10.1017/CBO9781139924146.002>
- Branch, R. M., & Varank, İ. (2009). *Instructional design: The ADDIE approach* (Vol. 722). Springer.
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of clinical epidemiology*, 46(5), 423–429. [https://doi.org/10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
- Chae, Y., & Davidson, T. (2026). Large language models for text classification: From zero-shot learning to instruction-tuning. *Sociological methods & research*, 55(2), 501–567. <https://doi.org/10.1177/00491241251325243>
- Chai, M., & Chen, J. (2025, 7–9 Nov. 2025). A Document Clustering Method with Classification Guidance and Sentence-BERT Model. 2025 5th International Conference on Digital Society and Intelligent Systems (DSInS), <https://doi.org/10.1109/DSInS68311.2025.11330016>
- Davila Delgado, J. M., Oyedele, L., Ajayi, A., Akanbi, L., Akinade, O., Bilal, M., & Owolabi, H. (2019). Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *Journal of Building Engineering*, 26, 100868. <https://doi.org/10.1016/j.jobbe.2019.100868>
- Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of clinical epidemiology*, 67(4), 401–409. <https://doi.org/10.1016/j.jclinepi.2013.12.002>
- Enke, J., Kraft, K., & Metternich, J. (2015). Competency-oriented design of learning modules. *Procedia Cirp*, 32, 7–12. <https://doi.org/10.1016/j.procir.2015.02.211>
- Frank, J. R., Mungroo, R., Ahmad, Y., Wang, M., De Rossi, S., & Horsley, T. (2010). Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach*, 32(8), 631–637. <https://doi.org/10.3109/0142159x.2010.500898>
- Hasson, F., & Keeney, S. (2011). Enhancing rigour in the Delphi technique research. *Technological Forecasting and Social Change*, 78(9), 1695–1704. <https://doi.org/10.1016/j.techfore.2011.04.005>
- Henri, M., Johnson, M. D., & Nepal, B. (2017). A Review of Competency-Based Learning: Tools, Assessments, and Recommendations. *Journal of Engineering Education*, 106(4), 607–638. <https://doi.org/https://doi.org/10.1002/jee.20180>
- Jünger, S., Payne, S. A., Brine, J., Radbruch, L., & Brearley, S. G. (2017). Guidance on Conducting and REporting DELphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliative medicine*, 31(8), 684–706. <https://doi.org/10.1177/02692163176906>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Kumar, S., Datta, S., Singh, V., Datta, D., Singh, S. K., & Sharma, R. (2024). Applications, Challenges, and Future Directions of Human-in-the-Loop Learning. *IEEE Access*, 12, 75735–75760. <https://doi.org/10.1109/ACCESS.2024.3401547>
- Levina, N., & Vaast, E. (2005). The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems¹. *Management Information Systems Quarterly*, 29(2),

335–364. <https://doi.org/10.2307/25148682>

- Liang, C.-J., Wang, X., Kamat, V. R., & Menassa, C. C. (2021). Human–robot collaboration in construction: Classification and research trends. *Journal of Construction Engineering and Management*, 147(10), 03121006. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002154](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002154)
- Mager, R. F., & Peatt, N. (1962). *Preparing instructional objectives* (Vol. 962). Fearon Publishers Palo Alto, California.
- Mulder, M. (2014). Conceptions of Professional Competence. In S. Billett, C. Harteis, & H. Gruber (Eds.), *International Handbook of Research in Professional and Practice-based Learning* (pp. 107–137). Springer Netherlands. https://doi.org/10.1007/978-94-017-8902-8_5
- Oh Heung, J., Chang, S., & Ashuri, B. (2023). Patterns of Skill Sets for Multiskilled Laborers Based on Construction Job Advertisements Using Web Scraping and Text Analytics. *Journal of Management in Engineering*, 39(3), 04023009. <https://doi.org/10.1061/JMENEA.MEENG-5243>
- Ojha, A., Zhang, Y., Shayesteh, S., Sharifironizi, A., Jebelli, H., & Akanmu, A. (2025). Advancing Construction Robotics Learning through Virtual Reality: A User Experience and Usability Study. In *Computing in Civil Engineering 2024* (pp. 110–119). <https://doi.org/10.1061/9780784486139.012>
- Olukanni, E., Akanmu, A., & Jebelli, H. (2025). Industry perception of competencies for human–robot collaboration in the construction industry: A Delphi study. *Frontiers of Engineering Management*, 12(4), 854–879. <https://doi.org/10.1007/s42524-025-4224-x>
- Olukanni, E., Akanmu, A., & Jebelli, H. (2026a). *Dataset associated with "A Data-Driven and Theory-Guided Framework for Human–Robot Collaboration Workforce Training in Construction Engineering and Management"* (University Libraries, Virginia Tech. <https://doi.org/10.7294/32394888>
- Olukanni, E., Akanmu, A., & Jebelli, H. (2026b). Systematic Literature Review of Competencies for Human–Robot Collaboration in Construction: A Task-Technology Fit Perspective. *Journal of Computing in Civil Engineering*, 40(1), 03125001. <https://doi.org/10.1061/JCCEE5.CPENG-6674>
- Pan, W., Chen, L., & Zhan, W. (2021). Positioning construction workers' vocational training of Guangdong in the global political-economic spectrum of skill formation. *Engineering, Construction and Architectural Management*, 28(9), 2489–2515. <https://doi.org/10.1108/ECAM-11-2019-0641>
- Perera, S., Babatunde, S. O., Pearson, J., & Ekundayo, D. (2017). Professional competency-based analysis of continuing tensions between education and training in higher education. *Higher Education, Skills and Work-based Learning*, 7(1), 92–111. <https://doi.org/10.1108/HESWBL-04-2016-0022>
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*. <https://doi.org/10.48550/arXiv.1908.10084>
- Reiser, R. A., & Dempsey, J. V. (2012). *Trends and issues in instructional design and technology*. Pearson Boston.
- Romão, A. C., Ribeiro, F., Sousa, L. M., & Neves, A. J. (2025). Mapping of Educational Course Descriptions to ESCO Competences Using Large Language Models. 2025 IEEE Global Engineering Education Conference (EDUCON), <https://doi.org/10.1109/EDUCON62633.2025.11016399>
- Saleema, J., Iyer, L. S., & Sankaran, L. (2025). Framework for Bridging the Skill Gap between Higher Education Institutions and Industry: A Machine Learning Approach. In *Cloud Computing for Smart Education and Collaborative Learning* (pp. 222–244). Chapman and Hall/CRC.
- Surana, S., Pathak, K., Gagnani, M., Shrivastava, V., M. T. R., & G, S. M. (2022, 16–18 March 2022). Text Extraction and Detection from Images using Machine Learning Techniques: A Research Review. 2022 International Conference on Electronics and Renewable Systems (ICEARS), <https://doi.org/10.1109/ICEARS53579.2022.9752274>
- Wu, S., Walzer, A. N., Kahlert, A., Dillenburger, B., & Hall, D. M. (2024). Understanding stakeholders' intention to use construction robots: a fuzzy-set qualitative comparative analysis. *Construction Robotics*, 8(1), 5. <https://doi.org/10.1007/s41693-024-00115-z>

- Zhang, M., Xu, R., Wu, H., Pan, J., & Luo, X. (2023). Human–robot collaboration for on-site construction. *Automation in Construction*, 150, 104812. <https://doi.org/10.1016/j.autcon.2023.104812>
- Zheng, J., Wen, Q., & Qiang, M. (2020). Understanding Demand for Project Manager Competences in the Construction Industry: Data Mining Approach. *Journal of Construction Engineering and Management*, 146(8), 04020083. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001865](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001865)