frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 29 January 2026
DOI 10.3389/fcomp.2025.1701666

OPEN ACCESS

EDITED BY
Jiajia Zhao,
Hanyang University, Republic of Korea

REVIEWED BY
Sheng Yang,
University of Guelph, Canada
Qian Bao,
Hanyang University, Republic of Korea

*CORRESPONDENCE
Pedro Ponce
✉ pedro.ponce@tec.mx

RECEIVED 08 September 2025
REVISED 09 December 2025
ACCEPTED 15 December 2025
PUBLISHED 29 January 2026

CITATION
Ponce P, Anthony B, Méndez JI, Bradley R, Guerrero J and Mata O (2026) Large language models and AI-driven virtual laboratory for FrED and FrED factory: materials, products, and sustainable digital manufacturing. *Front. Comput. Sci.* 7:1701666.
doi: 10.3389/fcomp.2025.1701666

Large language models and AI-driven virtual laboratory for FrED and FrED factory: materials, products, and sustainable digital manufacturing

Pedro Ponce[1]*, Brian Anthony[2], Juana Isabel Méndez[3], Russel Bradley[2], Joan Guerrero[1] and Omar Mata[1]

[1]Institute of Advanced Materials and Sustainable Manufacturing, Tecnologico de Monterrey, Mexico City, Mexico, [2]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States, [3]School of Architecture, Art and Design, Tecnologico de Monterrey, Monterrey, Mexico

Developing new products requires innovative materials and advanced manufacturing methods. Consequently, establishing specialized laboratories capable of producing new products or enhancing manufacturing processes has become essential. Additionally, the complexity of product design, which involves multiple subsystems, requires extensive iteration, making the process both challenging and costly. Evaluating manufacturing conditions further adds to these difficulties and expenses. In response, cutting-edge laboratories utilizing advanced technologies have been developed. These laboratories offer several advantages, such as remote operation, where equipment can be controlled and tests conducted systematically from a distance. Moreover, Virtual Reality (VR) laboratories have gained traction due to their lower costs and flexibility. VR laboratories can be adjusted and used to train students and operators through immersive technologies that simulate real-world scenarios. This paper proposes an innovative virtual laboratory deployed on Oculus Quest 3 and Android devices. The VR laboratory interacts with users through large language models. The VR laboratory features a virtual Fiber Extrusion Device (FrED) developed at MIT, as well as expert avatars specializing in specific topics, offering solutions to develop soft skills. Furthermore, the VR experience is tailored to the user's personality, enhancing the overall experience. Factory conditions are also simulated and optimized within the immersive laboratory using advanced optimization algorithms.

KEYWORDS

AI-driven sustainable manufacturing, LLM, soft skills, tailored VR lab, virtual model

Introduction

In recent years, technological education has undergone significant advances, making learning more interactive, dynamic, and realistic. This work focuses on utilizing technologies such as Virtual Reality (VR), Artificial Intelligence (AI), and immersive laboratories to enhance learning in engineering and manufacturing education. Integrating VR laboratories with AI-powered tutors and immersive simulations creates unique opportunities for enhancing practical skills and acquiring knowledge in a controlled environment.

Virtual Reality (VR) has gained recognition in STEM education, particularly for providing an immersive, hands-on learning experience that eliminates the need for physical equipment, making it ideal for high-risk and costly experimentation (Liu et al., 2024; Acevedo et al., 2024). However, challenges such as cost, accessibility, and the need for well-designed immersive experiences aligned with educational objectives persist (Acevedo et al., 2024). VR-based education can enhance learning outcomes, motivation, and user engagement, with a focus on thoughtful design to maximize its benefits (Acevedo et al., 2024; Yang et al., 2024).

VR integration in manufacturing processes has also advanced significantly. VR labs provide environments for interacting with virtual machines, allowing for training and experimentation without the costs or risks associated with real-world operations. Studies have shown the effectiveness of VR in training operators by replicating real-world scenarios, which enhances engagement and retention of skills (Hoang et al., 2022). VR laboratories also offer flexible learning and testing, allowing users to operate virtual machines remotely. Recent improvements focus on enhancing realism and responsiveness through the use of real-time physics engines and high-quality models (Oliveira et al., 2007).

AI-based tutoring systems ensure personalized, adaptive learning experiences catering to individual needs (Chen et al., 2020). Intelligent Tutoring Systems (ITS) provide 24/7 support and effectively complement standard teaching practices (Bezanson et al., 2023; Kim and Kim, 2020; Fazlollahi et al., 2022). AI tutors, as described by Fazlollahi et al., match the effectiveness of expert instruction, providing personalized learning pathways that would otherwise be unavailable (Fazlollahi et al., 2022). These AI systems have also been integrated into virtual learning environments to provide tailored content and enhance skill acquisition in various fields, including software engineering (Frankford et al., 2024). Large Language Models (LLMs), such as OpenAI's GPT, enhance user interactions in virtual learning by providing expert guidance, answering questions, and tailoring content to user needs (Pester et al., 2024). In manufacturing-oriented VR environments, LLMs can serve as virtual consultants, providing expertise in machinery operation, maintenance, and troubleshooting, while integrating domain-specific knowledge for industry-specific tasks (Khelifi and Morris, 2024).

On the other hand, advanced manufacturing technologies, including additive manufacturing, automated assembly, and digital twins, are transforming product development and production lines. Besides, Digital Twins offer real-time virtual replicas of physical systems, enabling the monitoring and optimization of manufacturing processes (Li, 2022). These innovations enhance predictive maintenance, quality control, and resource management, leading to increased productivity and lower costs. Combining VR labs with digital twins creates opportunities for process optimization and testing in a risk-free virtual environment (Li, 2022). Integrating Digital Twins in VR labs also bridges the gap between traditional hands-on and remote experiences, offering a more comprehensive learning environment (Alsaleh et al., 2022).

In addition, optimization algorithms play a crucial role in enhancing the performance of manufacturing systems in virtual environments. Metaheuristic algorithms, such as Genetic Algorithms, Particle Swarm Optimization, and Simulated Annealing, are used to optimize factory parameters, including resource allocation, energy consumption, and throughput (Hamid et al., 2014). These algorithms, implemented in VR labs, provide insights into optimal processes in a controlled, immersive setting, resulting in improved efficiency and reduced waste (Hamid et al., 2014).

Tailored gamification within educational systems helps improve motivation, engagement, and learning outcomes. Personalized, gamified systems adapted to learner preferences significantly enhance educational experiences (Oliveira et al., 2023). However, challenges include tailoring to individual styles and demonstrating impact (Oliveira et al., 2023). Personalizing educational experiences using psychological models, such as the Big Five Personality Traits, enables content and interaction to be tailored for individual users, resulting in increased engagement and performance (Chen et al., 2024). In VR labs, personality tailoring adjusts task complexity, pacing, and feedback tone, providing a more effective learning experience that considers individual differences (Chen et al., 2024).

As a result, recent studies have highlighted the importance of integrating soft skills training into STEM education. Soft skills, such as communication, teamwork, and leadership, are crucial for professional success but are often underrepresented in technical training. VR-based simulations have shown promise in teaching soft skills effectively (Hickman and Akdere, 2017; Abdelouahab, 2020; Caeiro-Rodríguez et al., 2021). VR-assisted environments offer the advantage of training both soft and hard skills, bridging the gap between theory and application (Abdelouahab, 2020). Virtual tutors can enhance decision-making and leadership skills by simulating real-world industry scenarios, thereby improving the transfer of skills to professional settings (Botke et al., 2018). Table 1 presents a comparison of educational technologies to provide a general overview.

Early work on virtual laboratories in engineering education established both the promise and the practical questions that later DT

TABLE 1 Comparison of technologies for education.

Technology	Advantages	Disadvantages
Virtual reality (VR)	Immersive learning enhances practical skills, ideal for complex subjects	High cost, requires specialized hardware, accessibility issues (Liu et al., 2024; Acevedo et al., 2024)
AI tutoring system (ITS)	Personalized learning, 24/7 availability, adaptive to student needs	Limited to certain subjects, requires careful design of feedback (Bezanson et al., 2023; Kim and Kim, 2020; Fazlollahi et al., 2022)
Gamification	Increases motivation and engagement, enhances the enjoyment of learning	Needs personalization, impact on learning outcomes not always clear (Oliveira et al., 2023)
AI-powered VR labs	Combines the benefits of VR and ITS, real-time adaptive learning	High development costs, complexity in integrating AI and VR effectively (Yang et al., 2024; Bezanson et al., 2023)
Soft skills training	Develops communication, leadership, and teamwork skills, with immersive practice in a safe setting	Intangible outcomes, difficult to evaluate skill transfer (Hickman and Akdere, 2017; Abdelouahab, 2020; Caeiro-Rodríguez et al., 2021)

and AI systems would inherit. In chemical and biochemical engineering, ViRILE provided a plant-scale simulator grounded in first-principles equations, raising enduring issues about fidelity, verification, and how closely simulations reflect real operating envelopes (Schofield, 2012). The next wave extended immersion to unit operations training; for example, an immersive crude-distillation experience showed that VR could scaffold procedure rehearsal and hazard awareness at the plant scale (Pirola et al., 2020). By 2021, reviews in chemical/biochemical engineering consolidated evidence that VR improves access to complex processes while highlighting gaps in assessment rigor and the cost/effort of content development (Kumar et al., 2021). Parallel surveys conducted in 2023 broadened the lens to encompass manufacturing and engineering education, cataloging VR/XR deployments and identifying persistent barriers, particularly in evaluation design and limited adoption beyond pilots (de Giorgio et al., 2023; Lampropoulos et al., 2025).

From 2024 onward, digital-twin exemplars demonstrated that fidelity and assessment can coexist. A VR bioreactor DT demonstrated realistic operator training, including infrequent event rehearsal, while documenting design principles, learning analytics, and evaluation methods that translate directly to discrete manufacturing or training contexts (Hassan et al., 2024). In 2025, comprehensive reviews mapped how DTs and virtual learning environments can support smart-manufacturing education, emphasizing human-centric, AI-enabled learning flows and multi-mode access (Filho and Junior et al., 2025). Scoping reviews in manufacturing and broader engineering education synthesized more than a hundred studies, offering taxonomies (domains, levels, entities) and practical roadmaps for adoption at scale while reiterating the need for stronger causal evaluation (Ipsita et al., 2025; Karanam and Hartman, 2025; Lampropoulos et al., 2025).

Concurrently, immersive AI tutoring matured from concept to reference architecture. An LLM-centric design for intelligent tutoring within VR specified core capabilities—real-time dialogue with non-player experts, hand/gaze/haptic multimodality, as well as synchronized speech and embodiment—creating a concrete blueprint for avatar-mediated instruction in labs (El Hajji et al., 2025). The broader ITS literature in 2025 introduced methodological guardrails, with systematic reviews generally documenting positive learning effects but calling for stronger experimental designs (e.g., blocking, pre/post ANCOVA) and clearer links between telemetry and outcomes, as well as an AI-driven intelligent tutoring system (Létourneau et al., 2025; Liu et al., 2025). At the authoring layer, studies in Nature's education journals have shown that lesson plans can be improved through LLM-simulated teacher–student interactions and structured human–LLM workflows, which are useful for designing avatar playbooks, prompt governance, and reflection loops in immersive labs (Feisel and Rosa, 2005; Flores Romero et al., 2025). Finally, open-source XR agent platforms lowered the barrier to deploying LLM-driven virtual humans in Unity-based environments, accelerating reproducibility and comparative evaluation (Shoa and Friedman, 2025).

Taken together, the field has converged on three actionable insights for manufacturing education. First, DT-backed VR can deliver authentic process and factory experiences when fidelity is paired with explicit assessment scaffolds (Hassan et al., 2024; Junior et al., 2025; Karanam and Hartman, 2025; Peterson et al., 2025). Second, LLM-driven avatars are no longer speculative; reference architectures and toolkits exist, but evaluation must tie dialogue and behavior telemetry to validated learning outcomes using robust designs (El Hajji et al., 2025; Létourneau et al., 2025). Third, new authoring paradigms, including LLM-based simulation, structured prompting, and open XR agents, enable scalable, multimodal laboratories, provided governance (safety, provenance, and bias) and accessibility (multi-device and multi-language) are treated as first-class design constraints (Feisel and Rosa, 2005; Ipsita et al., 2025; Shoa and Friedman, 2025).

Despite rapid progress within each strand of research, the literature still lacks a truly end-to-end framework that integrates immersive VR, LLM-driven tutoring or avatars, and operational Digital Twins and Virtual Models into a single, assessable learning environment for sustainable manufacturing, while simultaneously cultivating soft skills such as communication, ethical reasoning, teamwork, and decision-making. Prior studies typically examine these components in isolation, for example, VR labs without causal evaluation, DT implementations without explicit pedagogy, or human–AI workflows detached from process fidelity. As a result, key questions remain open regarding the alignment between fidelity and assessment, human-in-the-loop safety and governance, multi-device and multi-language accessibility, and the linkage between system telemetry and validated learning outcomes.

This article addresses that gap by presenting a unified VR–LLM–VM (virtual model) laboratory that combines realistic factory and process behavior with LLM-guided avatar guidance and analytics-supported decision tools (e.g., optimization and forecasting), while embedding soft-skills training through avatar-mediated critiques, oral defenses, and ethical audit trails. The environment is engineered for inclusivity (headset and mobile pathways, bilingual interaction), measurement (rubrics tied to system telemetry and key performance indicators), and governance (traceable prompts, bounded actions, rollback), and it is paired with a rigorous evaluation design that uses pre/post testing, blocked assignment, and blinded rating to generate credible evidence at scale. Although VR has been used effectively in STEM education, with a primary emphasis on procedural and technical competence (Liu et al., 2024; Acevedo et al., 2024), the systematic development of soft skills in immersive settings remains comparatively underexplored (Hickman and Akdere, 2017; Abdelouahab, 2020; Caeiro-Rodríguez et al., 2021). Existing studies indicate that immersion and realism can enhance soft-skill rehearsal and near transfer; however, consistent far transfer to authentic workplaces is challenging because these skills are partly tacit, highly contextual, and difficult to measure objectively at scale (Caeiro-Rodríguez et al., 2021). A promising direction is the integration of LLM-based virtual tutors that adopt varied, role-specific personae (for example, supervisor, peer, customer), creating context-specific dialogues, critiques, and decision checkpoints that mirror real interactions and yield structured, auditable evidence for assessment (Botke et al., 2018). Embedding these LLM roles directly into VR task flows, alongside high-fidelity virtual models of the extruder and factory (with associated production metrics), enables concurrent practice of both technical operations and soft skills without yet requiring a full real-time digital twin connection. Telemetry such as dialogue turns, quality of rationale, and team-coordination traces can be mapped to rubrics and performance outcomes, thereby addressing both the learning-design gap and the evaluation challenge documented in prior work (Liu et al., 2024; Acevedo et al., 2024;

Hickman and Akdere, 2017; Abdelouahab, 2020; Caeiro-Rodríguez et al., 2021; Botke et al., 2018).

As a result, this combination of technological tools enhances technical skill training while developing critical soft skills, offering personalized content and continuous, contextual feedback. On the other hand, the use of a didactic Fiber Extrusion Device (FrED) offers an interactive, hands-on experience that enables trainees to engage directly with manufacturing processes, thereby enhancing their understanding of sustainable digital manufacturing practices.

It is essential to consider that implementing a virtual reality course laboratory presents several recurring limitations. Physiological effects such as dizziness, nausea, headache, eye strain, and difficulty concentrating can restrict session length and exclude a subset of learners. Rich visuals, fast motion, and nonessential interactables often lead to cognitive overload and distraction, diverting attention away from the learning objective. Ergonomic issues, such as fatigue, neck strain, and heat buildup, accumulate during extended headset use, forcing shorter, more frequent sessions that reduce instructional time. Accessibility remains uneven: fully immersive head-mounted displays are a poor fit for students with visual impairments and challenging for some neurodivergent or vestibular-sensitive learners, making it hard to guarantee true equivalence of experience. Safety and liability concerns persist because occluded vision and cabling introduce collision and trip hazards, requiring supervision, clear space, sanitation routines, and incident procedures (Soliman et al., 2021). In addition, cost and resources are substantial: high-fidelity headsets, capable computers, tracking systems, and consumables demand ongoing funding, while maintenance and replacement cycles are accelerated by wear and rapid obsolescence. Infrastructure and logistics complicate delivery; VR labs require generous floor space, reliable power and ventilation, storage, booking systems, and device fleet management. Throughput is limited by the number of headsets and the time required for the turnaround between groups. Development is complex, as educational scenarios require expertise in game engines, three-dimensional modeling, interaction design, and software engineering. Commercial assets tend to prioritize entertainment over assessment or pedagogy, so bespoke scripting is common (Soliman et al., 2021).

Pedagogical alignment and validation are non-trivial: without an explicit mapping to learning outcomes and iterative testing, engagement may increase without corresponding measurable learning gains. Assessment and data integrity pose further challenges because telemetry must be interpreted with caution, and proctoring within immersive environments is particularly difficult. Privacy and ethics introduce administrative burdens, as VR can capture sensitive signals such as gaze, posture, voice, and biometrics that require strict consent, minimization, retention, and access controls. Staffing and training needs increase for instructors and assistants who must handle setup, troubleshooting, hygiene, and safety, and reliability remains fragile due to firmware updates, tracking glitches, and driver conflicts that can derail scheduled sessions. Interoperability with campus platforms is often fragile, resulting in *ad hoc* integration with learning-management systems and analytics. Finally, transfer to real equipment is not guaranteed; gaps in haptics, force feedback, and material behavior mean that skills learned in virtual environments may not generalize without complementary hands-on work (Soliman et al., 2021). As a result, this paper attempts to address some of these limitations by utilizing an AI-enabled virtual laboratory that integrates large language models with the FrED extrusion system within an interactive VR environment. In this environment, learners interact in real-time with expert avatars and other non-player characters, which maintain their own conversation histories to preserve context. To broaden access, the lab is deployed in two complementary formats: a fully immersive Oculus Quest 3 experience with 3D avatars, and a performance-optimized Android version that omits avatars to ensure smooth operation while preserving core learning goals. Both versions support voice and text interaction in English and Spanish, enhancing inclusivity without compromising functionality.

Across platforms, the laboratory offers a coherent suite of 14 interactive elements that guide learners from profile setup and subsystem exploration to AI-assisted analysis. Within this progression, predictive analytics play a central role. For production optimization, a loss-driven genetic algorithm iteratively searches operating configurations using tournament selection, crossover, and mutation. The objective function penalizes low throughput, oversized buffers, and failure or repair rates that deviate from baseline targets, guiding the search toward efficient and cost-effective operating points and demonstrating convergence across successive generations. In parallel, a neural network trained on data from VR-simulated robotic arms forecasts line status, enabling proactive decisions about operations and maintenance inside the factory scenario. The technical content is framed by a materials-education perspective through the Penta-S framework, which emphasizes smart, sustainable, social, sensing, and safe considerations. This lens connects process choices to product properties and sustainability outcomes, ensuring that optimization and forecasting are not purely algorithmic but also pedagogically meaningful. Taken together, the design fosters critical thinking, problem-solving, collaboration, adaptability, and leadership, while remaining inclusive and portable across various hardware tiers.

What is FrED?

FrED (fiber extrusion device) is a compact, affordable desktop fiber extrusion system designed primarily for educational purposes, developed at MIT. While not a high-fidelity replica of industrial setups, it simulates fiber draw mechanisms, enabling hands-on learning in smart manufacturing, control systems, data acquisition, and computer vision. Acting as a bridge between theory and practice, FrED enhances understanding of advanced manufacturing concepts. FrED consists of key components: an extrusion subassembly, a diameter measurement subassembly, a cooling subassembly, and a spooling subassembly. The process involves heating and melting material in the extrusion subassembly, measuring fiber diameter with a USB camera, cooling the fiber, and winding it evenly in the spooling subassembly. A built-in control feedback loop supports real-world learning in manufacturing environments. Developed to support MIT courses, FrED has evolved through multiple iterations. The first "Research FrED" targeted high-performance educational use (Bradley, 2023). One of the first versions prioritized affordability with simplified designs and lower-cost materials, making it widely accessible. This version was then enhanced with cooling, control algorithms, and mechanical stability to achieve even lower production costs (Xu, 2024). The next version, detailed by Rosko (2024) and Zhang (2024), introduced closed-loop control for precise diameter regulation, enhanced mechanical stability, an optimized cooling system, and a redesigned user interface. These upgrades represent a significant advancement, featuring dynamic motor speed and heating adjustments based on real-time feedback, which further bridge the gap between academic learning and industrial practices. Figure 1 illustrates FrED's most recent version, bringing advancements in cost efficiency, mechanical stability, control systems, and educational usability. FrED has the following main systems:

a) Extrusion sub-assembly that extrudes the preform using a stepper motor that pushes the preform into the heater block with a gear. The heater block heats up the preform to form a fiber.
b) Fiber cooling sub-assembly that has two fans that cool the fiber before it reaches the pulleys of the diameter measurement.
c) The diameter measurement sub-assembly consists of a pulley system that maintains the position of the fiber for a USB camera to take diameter measurements optically of the fiber.
d) The spooling sub-assembly has a removable spool that is controlled by a DC motor and gearbox to collect extruded fiber neatly through rotation and single-axis translation.
e) A PCB board and Raspberry Pi controller stored below the base plate. These control the FrED.

With these enhancements, FrED is well-suited as a training tool in virtual reality (VR). Its compact design, real-time control, and data acquisition features enable immersive simulations of the fiber extrusion process. This setup allows learners to safely practice skills in a realistic, controlled environment, deepening their understanding of complex manufacturing concepts through hands-on interaction.

Methodology

This paper proposes the development of an innovative virtual reality (VR) laboratory using virtual models that integrates Large Language Models (LLMs) to enhance the understanding of subsystems in the Educational Extruder (FrED) and its associated factory. A notable feature of FrED is its ability to assess materials, providing insights into the feasibility of material substitution for improved performance. Figure 2 illustrates the flow diagram of the proposed VR laboratory, highlighting its phases and components where optimization algorithms and AI could be implemented (Ponce and Ponce, 2011).

The VR laboratory offers a personalized educational experience by tailoring the learning environment to the Big Five personality traits. Tailoring the environment to individual profiles ensures that educational content resonates with each learner's preferences. Gamification further enhances this personalized approach by tailoring rewards to individual learner traits, thereby fostering motivation and engagement.

The core interaction phase bridges the gap between theory and practice, enabling users to engage with specific extruder subsystems within the VR environment, including control mechanisms, power

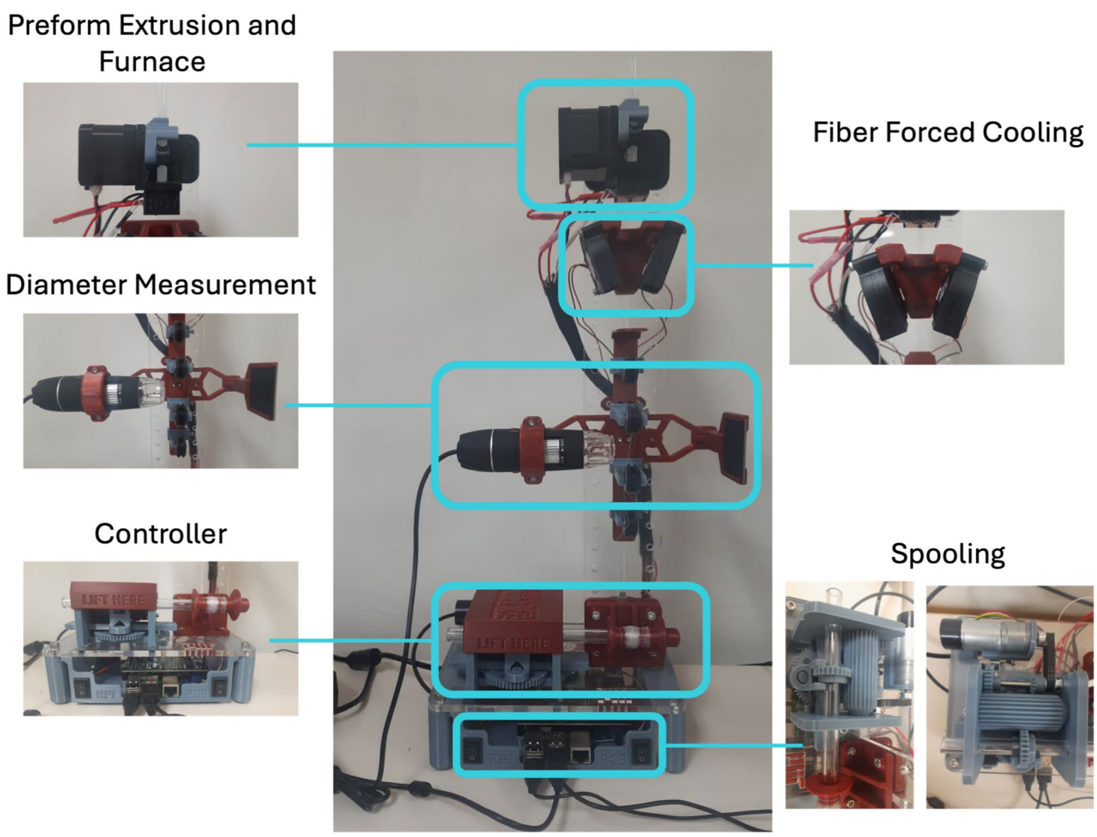

FIGURE 1
FrED 2024 version.

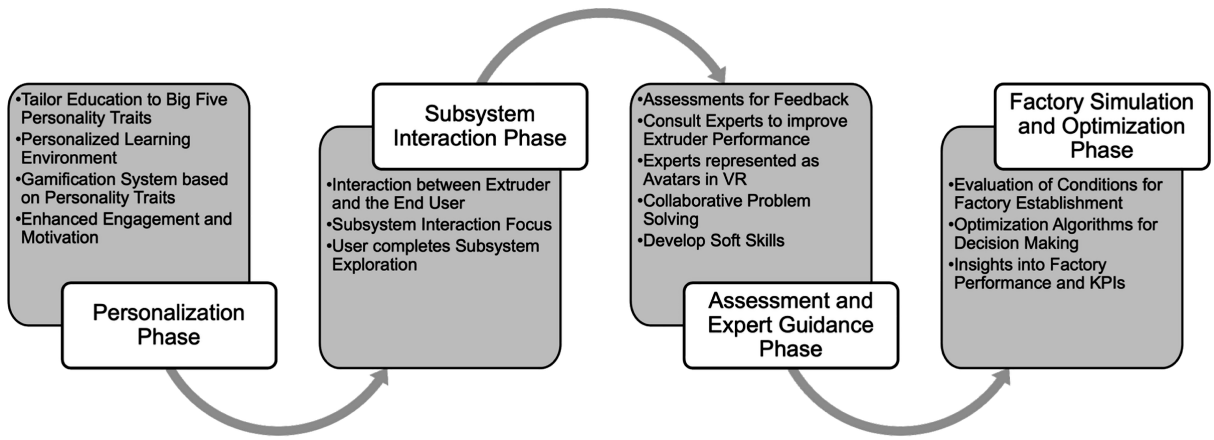

FIGURE 2
Flow diagram of the proposed laboratory, divided into four steps: personalization, subsystem interaction, assessment and expert guidance, and the factory simulation and optimization phase.

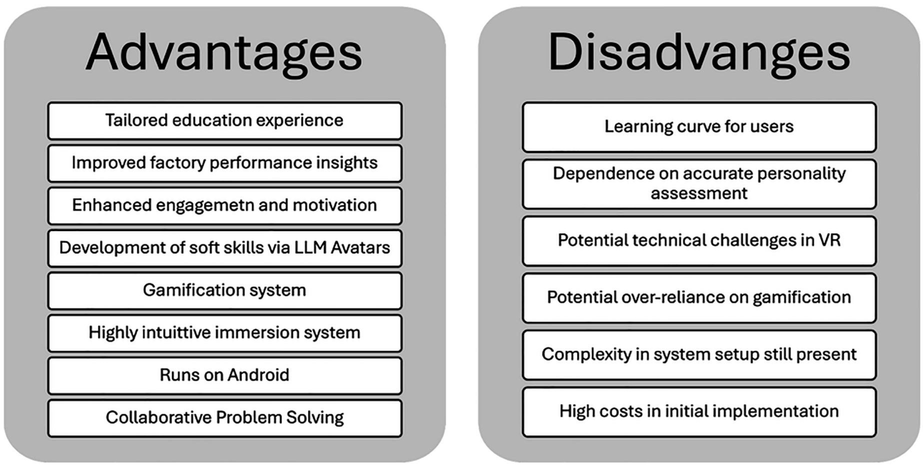

FIGURE 3
Main advantages and disadvantages of the proposal.

systems, and material evaluation tools. Each interaction offers an in-depth understanding, ensuring focused and effective learning. Users then complete assessments to identify misunderstandings, serving as feedback mechanisms to reinforce comprehension (Molina et al., 2024).

After assessments, users engage with experts represented as avatars in the VR environment, guiding them through complex challenges related to the extruder's operation. These experts assist users in resolving technical issues, promoting collaborative problem-solving, and developing essential soft skills, including teamwork and effective communication.

Another key aspect of the laboratory is its simulation of factory setup and operation. Users learn about practical considerations for establishing and optimizing a factory, utilizing advanced optimization algorithms to enhance data-driven performance. These simulations offer practical insights into decision-making and strategy implementation, enabling the achievement of operational excellence (Ponce-Cruz et al., 2020).

The VR laboratory represents a dynamic platform for understanding manufacturing systems (Peniche et al., 2012; Rubio et al., 2005), integrating technical education with personal growth. By incorporating LLMs, personality-based customization, gamification, and optimization algorithms, the laboratory equips users to tackle real-world challenges with confidence.

It is worth noting that this laboratory can also run a lighter version on Android, enabling the educational proposal to be accessed on Android devices. This broadens the reach, enabling more users to benefit from the program. Figure 3 outlines the core advantages and potential disadvantages of this proposal.

General methodology description

In general, this study evaluates a unified VR–LLM–VM (virtual models) learning environment that begins with device-level mastery on the FrED educational extruder and then scales the same concepts to a factory/production-line context where students optimize flow, reliability, buffers, and cost using analytics (GA/ANN). The experiment compares learning and transfer across platforms (Oculus Quest 3 VR; Android) and interaction modes (with/without LLM avatars), using common tasks, rubrics, and telemetry to ensure comparability. The goal is to determine whether early, hands-on understanding of extruder physics, sensing, and multiloop control (Phase 1) improves the quality and speed of decisions when students face line-level trade-offs in throughput, WIP (work in progress), reliability, energy, and sustainability (Phase 3), with expert avatars (Phase 2) providing structured guidance and soft-skills practice along the way.

Proposed experimental setup:

- Participants & assignment: senior undergraduate/early graduate engineering students are block-randomized by a short prerequisite quiz and preferred device/language into four conditions: Full VR–LLM, VR-only, LLM-only (mobile/desktop), and business-as-usual control.
- Platforms: two builds deliver identical learning objectives: (1) Oculus Quest 3 (fully immersive, 3D avatars), (2) Android (performance-optimized, same tasks; avatars omitted).
- Core modules:

  o FrED device module (extruder anatomy, sensing, control, safety, test/characterization).
  o Avatar guidance module (diagnosis, safe parameter/policy edits, ethics & communication).
  o Factory module (production-line optimization via GA; line-status forecasting via ANN; Penta-S sustainability lens).

- Data & instruments: FrED rubric + CTS rubric, pre/post knowledge tests, capstone defense, and system telemetry (tool use, GA convergence, ANN metrics, dialogue traces).

In this proposed laboratory, the curriculum could be intentionally staged so that the variables students control on FrED, temperature zoning, torque/speed envelopes, sensor placement and calibration, controller tuning, and safe interlocks—map directly to the variables that drive line performance (failure/repair rates, buffer sizes, cycle times, energy per unit, defect risk). Students first learn to sense, control, and validate at the device level; then they encounter the same constructs as parameters in a production line, where they must balance throughput, WIP, reliability, and cost. This continuity prevents a "hard jump" from a single machine to a factory: the factory simply aggregates device-level physics and control decisions into system-level trade-offs that are optimized with GA and forecast with ANN.

Learning objectives (stated before procedures):

1. Device-level technical mastery (Phase 1): decompose the extruder into subsystems; derive and apply governing relations; instrument and control multiloop dynamics; run safe tests; generate clean, versioned datasets.
2. Human–AI collaboration & soft skills (Phase 2): obtain targeted diagnostics from avatars; translate natural-language guidance into safe, typed actions with bounds/units; design avatar-in-the-loop experiments; communicate decisions ethically with auditable logs.
3. Factory-level decision making (Phase 3): optimize buffers/reliability/throughput/cost with a loss-driven GA; anticipate states and maintenance with ANN; justify policies under the Penta-S sustainability frame; defend decisions to technical and executive audiences.

Minimum apparatus and materials required:

- Software/hardware: unity 5 application; Oculus XR plugin (VR); custom Android input; ChatGPT-based avatar services; Firebase for profiles/telemetry; in-app oscilloscope and CSV viewers; GA (tournament selection, crossover, mutation) and ANN (trained on VR-simulated robotic-arm data).
- Instructional assets: three activity tracks (Energy Efficiency Mastery; Precision & Quality Assurance; Stability & Control), avatar playbooks, and bilingual prompts (EN/ES).

Measures and evidence:

- Primary: capstone performance (FrED rubric + CTS composite) on a novel factory scenario with a 3-min oral defense.
- Secondary: knowledge/skill gain (pre–post), transfer efficiency (time-to-stable policy; constraint violations), decision quality (expert scores), inclusivity (device/language subgroups).
- Mechanism traces: GA convergence and final loss; ANN accuracy/confusion matrix; avatar dialogue density/intent; oscilloscope/CSV interaction logs.

The FrED VR laboratory is specifically designed to develop communication, teamwork, leadership, and informed decision-making in sustainable manufacturing scenarios. Radianti et al. (2020) reports that a small percentage of the surveyed immersive VR applications target these types of learning content, and that most evaluations focus on usability and user experience rather than on measurable learning or skill development. Against this backdrop, FrED VR Lab positions soft skills not as a secondary outcome but as a central target of the VR–LLM experience, embedded in realistic factory-level decision problems in engineering education.

Soft-skill outcomes are operationalized through an analytic rubric that combines a FrED-specific performance scale with a Creative Thinking Skills (CTS) rubric, reflecting the roles of critical thinking, problem-solving, collaboration, and leadership, which were previously identified as essential but often underrepresented in technical training. The composite rubric includes observable indicators such as: clarity, structure, and audience awareness when explaining factory-level policies; justification of decisions in terms of trade-offs among throughput, reliability, and sustainability; collaborative problem-solving (e.g., referencing teammates' ideas, negotiating constraints, and reaching shared decisions); and leadership behaviors, such as coordinating roles, framing next steps, and managing risk. These descriptors are applied to students' capstone performance on a novel factory scenario and to a three-minute oral defense, which trained evaluators independently rate. This approach directly addresses a challenge highlighted in prior work: soft-skill outcomes are often considered intangible, and their transfer in VR-based training is challenging to evaluate.

Beyond human ratings, the environment logs rich dialogue traces between learners and LLM-driven avatars (e.g., turn counts, dialogue acts, and quality of rationale) alongside in-world actions (policy edits, constraint violations, and experiment design choices). These traces are mapped to the rubric dimensions to provide a complementary, telemetry-based view of soft-skill practice—for example, how often learners proactively request feedback, how they negotiate conflicting objectives with the avatar, or how they revise decisions after critical questioning. In Paszkiewicz et al. (2021), it is shown that only a small fraction of existing studies exploit exams, expert judgments, or sensor/trace data to assess learning outcomes, despite frequent claims about the effectiveness of VR for skills development. By combining expert ratings with trace-based analytics, the FrED lab responds to this gap with a multi-method evaluation strategy tailored to soft skills in engineering.

To evaluate transferability, soft-skill scores and decision-quality indicators will be analyzed across conditions (full VR–LLM, VR-only, LLM-only, and business-as-usual) and across tasks that differ in proximity to the training context. Near transfer is assessed within the virtual factory using measures such as time-to-stable policy, constraint-respecting decisions, and expert ratings of justification quality. Far transfer is examined through follow-up oral defenses and written reports in related courses or projects, applying the same rubric to determine whether the communication, teamwork, and leadership behaviors practiced in the virtual lab reappear in new, non-VR tasks. This design directly addresses calls from the systematic review for more robust evaluation of learning outcomes, not only user experience, and for more substantial evidence on how soft skills trained in immersive VR transfer to authentic higher-education and professional contexts.

Tailoring education—requirements of operators

To enhance the learning experience, learners are profiled based on the Big Five Inventory (BFI) and categorized into gamified player types. The Openness trait appreciates divergent thinking, curiosity, and creativity. Conscientiousness indicates a rule-following attitude with clear goals. Extraversion involves social interaction and optimism. Agreeableness is associated with altruism and tolerance, while neuroticism is related to impulsiveness and stress.

Personality traits classification

An online survey with 35 questions was conducted from November 2020 to December 2022, collecting 645 responses. The dataset was filtered to include five questions for each personality trait and 30 questions on gamification preferences. The survey aimed to understand user preferences in three areas: reward types, main page elements, and activity interests within an educational platform.

(1) Reward preferences: users were asked, “What type of rewards would you prefer after completing exercises or activities?” Options included badges, coupons, no recognition, random rewards, sharing points, social recognition, money, extra points, store discounts, physical rewards, and diplomas.
(2) Main page preferences: users were asked, “What game elements would you prefer on the platform’s main page?” Options included avatar, total points, username, top 5, active challenges, improvement tips, winning tips, recent badge, pending challenges, leaderboard, and using their real name.
(3) Activity preferences: users were asked, “What types of activities interest you on an educational platform?” Options included activity badges, individual challenges, feedback, community challenges, social sharing, team challenges, topic-linked activities, and freedom to perform activities without affecting scores.

Responses were evaluated using a 5-point Likert scale to measure user preferences, ranging from “strongly disagree” (1) to “strongly agree” (5), which were then linked to different gamification options. A probabilistic approach was used to analyze results, with a threshold of 0.5 set to balance excluding unlikely options while capturing relevant elements (the highest probability was 0.75).

Figure 4 illustrates the game elements selected based on personality traits. The most preferred gamification elements included store discounts, diplomas, avatar, username, improvement tips, individual challenges, feedback, and topic-linked activities. Based on the results depicted in Figure 4, a tailored environment should consider the gamification elements depicted in Figure 5.

FrED’s rubric

In addition, the FrED’s rubric is a comprehensive tool for evaluating advanced manufacturing systems. It defines subsections such as digital systems, control systems, electric systems, and sensor actuators, with each subsystem having its own rubric. Figure 6 presents the rubric, which assesses extruder design across five key criteria: Design, Integration into Manufacturing Processes, Performance and Cost Efficiency, Innovation and Problem Solving, and Utilization of Technology and Tools. This image illustrates the logical structure and detailed considerations that ensure a comprehensive evaluation of extruder design projects, facilitating continuous improvement and alignment with industry standards.

This rubric provides a structured and fair assessment of extruder design projects, covering critical areas such as design quality, integration, performance, innovation, and technology utilization. Each criterion ensures that all aspects are thoroughly evaluated, leading to better outcomes and continuous improvement in extruder design.

FrED’S training activities and system division

FrED has multiple subsystems, including the Electronics Stage, Temperature Control System, Digital Control System, Computer Vision System, Mechanical Tensioning and Spooling System, and Sensor and Driver Integration, which address distinct aspects of the fiber extrusion process. The interconnected functionality of these systems highlights the importance of precise adjustments and real-time monitoring in ensuring high-quality fiber output.

To facilitate a structured, gamified learning experience, the training program is divided into three activities, depicted as challenge or activity elements within the VR environment (Figure 7), which reflect the distinct operational goals of each subsystem.

Each activity is designed to tackle a specific aspect of FrED’s operation, guiding learners through real-world scenarios that

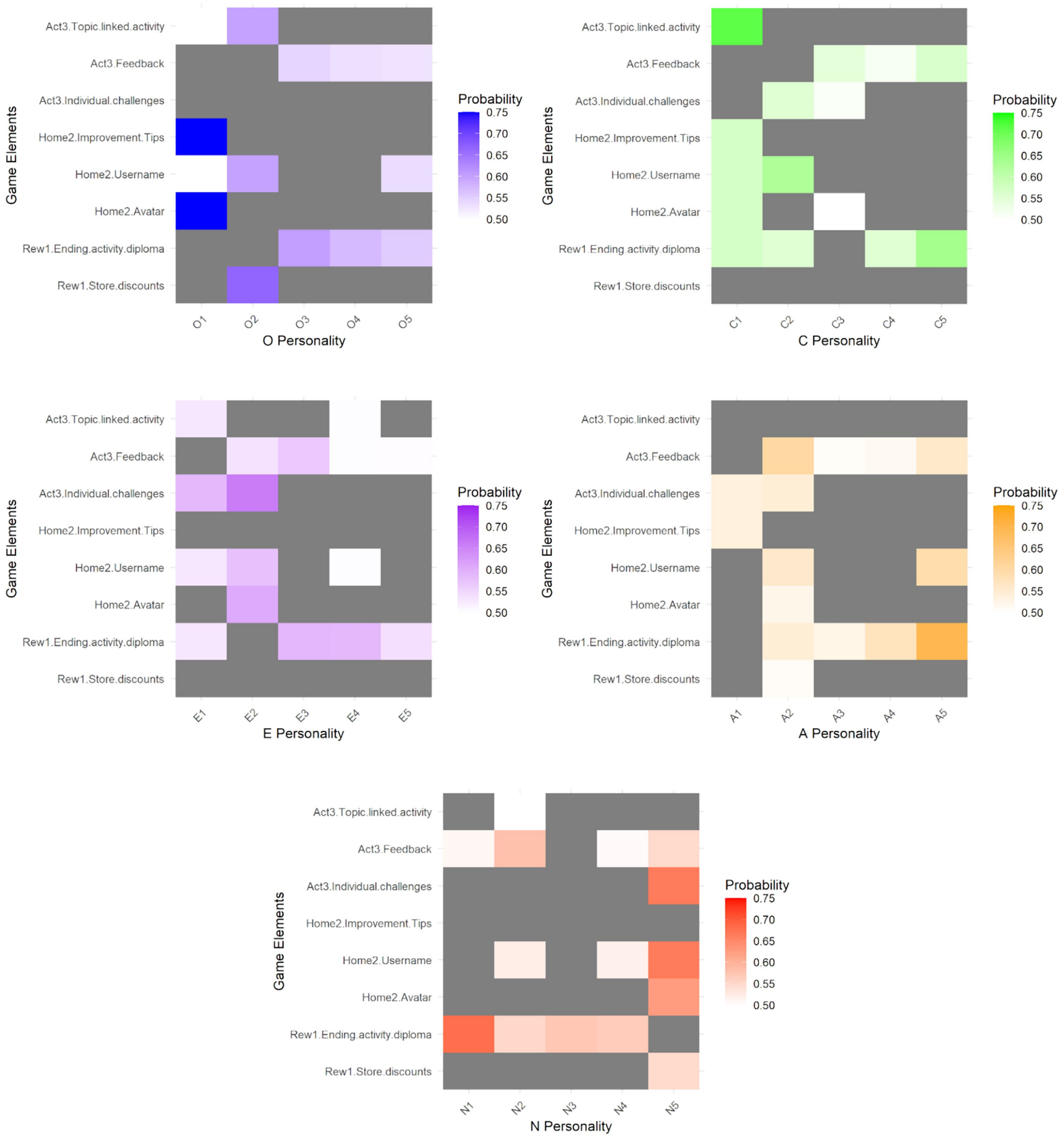

FIGURE 4
Game elements are based on the five personality traits: Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). Being 1, the strongly disagree option, to 5, from strongly agree, in Likert responses.

challenge them to apply technical knowledge in a practical context. The three activities are:

1) Activity 1: energy efficiency mastery: learners engage in scenarios that challenge them to optimize FrED's energy efficiency by adjusting settings in the Power Electronics Stage and Temperature Control System.
2) Activity 2: precision and quality assurance: this activity focuses on enhancing the quality and consistency of fiber production through precise control adjustments in the Computer Vision System and Digital Control System.
3) Activity 3: stability and control: the goal of this activity is to ensure system stability and responsiveness, focusing on preventing material breakage and lag in system responses through the integration of the Mechanical Tensioning and Spooling System with sensors and Drivers.

The scoring system across all activities is designed to encourage learners to improve and reinforce their knowledge. Each activity consists of four scenarios, with first-attempt correct answers earning 25% and second-attempt answers yielding 20%. Once the activity is completed, a diploma will appear with one of the

Trait	Likert scale 1	2	3	4	5
O	Avatar, Improvement Tips	Store discounts, Username, Topic linked with the activity	Ending activity diploma, Feedback	Ending activity diploma, Feedback	Ending activity diploma, Username, Feedback
C	Ending activity diploma, Avatar, Username, Improvement Tips, Topic linked to the activity	Ending activity diploma, Username, Feedback	Individual challenges	Ending activity diploma, Feedback	Ending activity diploma, Feedback
E	Ending activity diploma, Username, Individual challenges, Topic linked to the activity	Avatar, Username, Individual challenges, Feedback	Ending activity diploma, Feedback	Ending activity diploma, Username, Feedback, Topic linked to the activity	Ending activity diploma, Feedback
A	Improvement tips, Individual challenges	Store discounts, Ending activity diploma, Avatar, Username, Individual challenges, Feedback	Ending activity diploma, Feedback	Ending activity diploma, Feedback	Ending activity diploma, Username, Feedback
N	Ending activity diploma, Feedback	Ending activity diploma, Username, Feedback	Ending activity diploma	Ending activity diploma, Username, Feedback	Store discounts, Avatar, Username, Individual challenges, Feedback

FIGURE 5
5-Likert response to gamification by each personality trait, with 1 being strongly disagree and 5 being strongly agree. O = Ocean, C = Conscientiousness, E = Extraversion, A = Agreeableness, N = Neuroticism.

following seals depending on the student's knowledge level (see Figure 8).

Discovering Penta-S materials

Understanding the materials used in the manufacturing process is crucial to predicting the final product's features. Additionally, it is essential to outline how the material can be improved. The VR laboratory features a materials expert who can assist the end user in understanding the material and its properties, enabling the user to propose a new material for the application. Nowadays, a broad spectrum of materials can be used in manufacturing systems, so selecting the correct material for a learning person could be challenging because the person requires a holistic view of the materials. Additionally, the recycling process of materials must be considered to modify the product life cycle and ensure the best possible economic and ecological outcomes. The study of materials in manufacturing systems affects the machinery process, enabling increased productivity and quality control. The materials are studied under the concept of Penta-S materials. These materials include the following five features: smart, sustainable, social, sensing, and safe (Molina et al., 2025).

Tailored application—FrEd virtual laboratory

The virtual environment was developed using Unity Engine 5 and Oculus Quest technology, simulating the laboratory environment with key elements outlined in this study. The application is designed for two platforms: Oculus Quest 3 and Android devices.

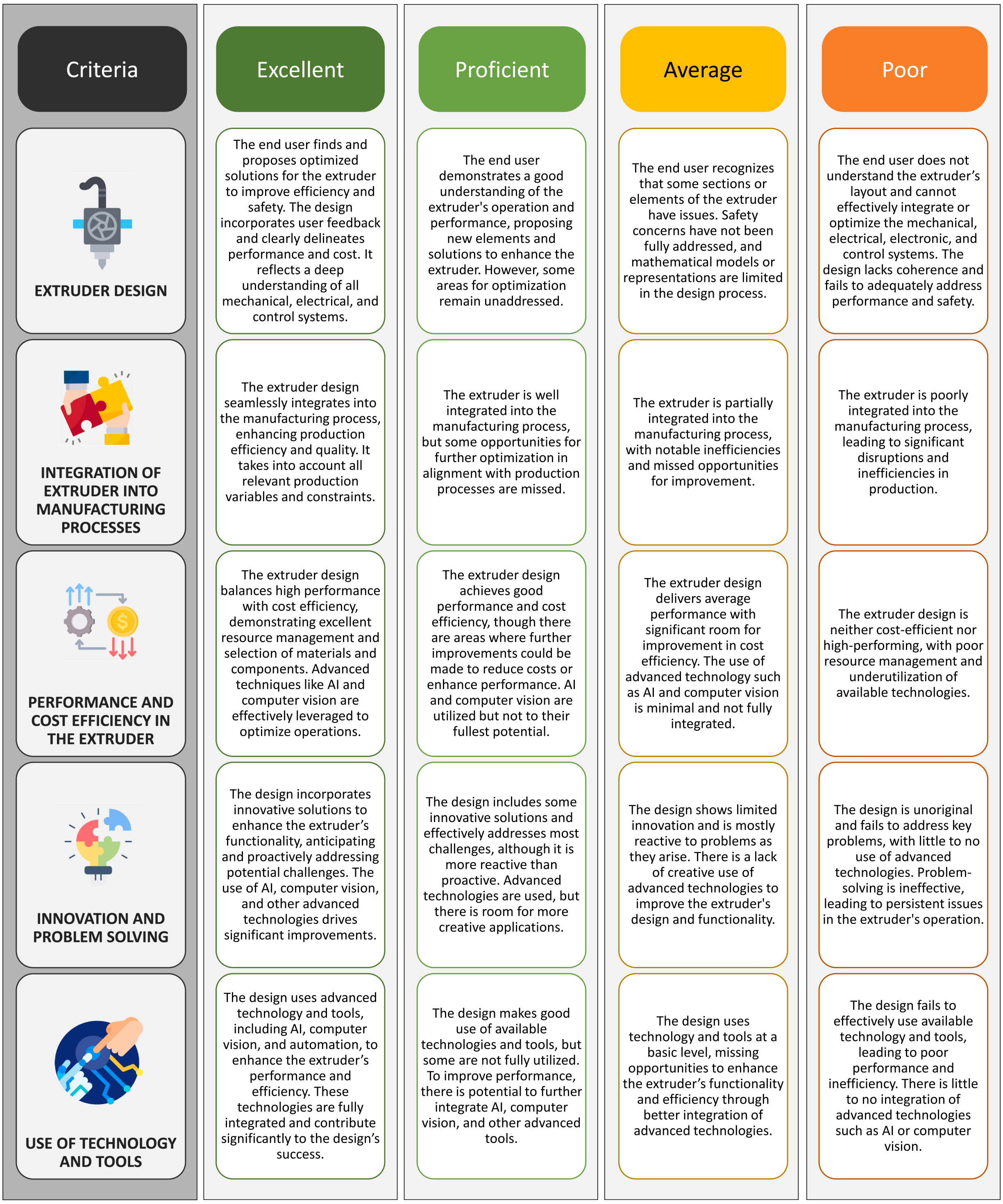

FIGURE 6
FrED's rubric showcases each criterion with the level of proficiency.

Key features include a player input controller tailored for both VR and Android versions. The integration of OpenAI allows interactive, natural language conversations with virtual experts, creating a personalized setup experience.

For the Oculus Quest 3, a customized version of the Oculus XR Plugin was utilized to handle player movement and interactions, with a focus on hand tracking, head movement, keyboard input, and object interaction, ensuring seamless VR engagement. For Android, a custom script enables 3D movement through on-screen joysticks and gestures, providing an intuitive interface like the VR version.

The saving system stores progress through local storage, with Firebase Realtime Database ensuring cloud backups. The application also includes interaction with the FrED machine, allowing users to visualize subsystems and use an integrated oscilloscope for real-time analysis of output data, offering an interactive way to observe and manipulate machine performance.

FIGURE 7
Activities presented in the VR environment. The seals are associated to each activity.

FIGURE 8
A diploma is sealed when the activity is completed.

Manufacturing proposal: soft skills in engineering

Problem-solving is a cornerstone of engineering, but it does not stand alone—it is closely intertwined with skills such as innovation, brainstorming, critical thinking, research, and leadership (Caeiro-Rodríguez et al., 2021). These interconnected skills are essential not only for solving technical challenges but also for enhancing project management and business efficiency.

Figure 9 illustrates key soft skills in engineering, including critical thinking, innovation, problem-solving, and teamwork. Several important branches are highlighted:

(1) Brainstorming: linked to critical thinking and innovation, demonstrating their role in effective brainstorming sessions in engineering.
(2) Problem solving: involves skills like validation, research, solution development, and problem identification, all crucial to the problem-solving process.
(3) Team collaboration: emphasizes working together effectively, supported by skills like documentation and validation. Collaboration requires communication and verifying outcomes.
(4) Leadership: connected to critical thinking and innovation, indicating that strong leadership in engineering requires both creative and analytical abilities.

How to implement the proposed laboratory in an engineering class

Traditional engineering instruction is predominantly lecture-based, with information transmitted via slides, boards, and instructor exposition (Soliman et al., 2021). This mode of delivery tends to privilege instructor activity over student agency, which can limit opportunities for inquiry and knowledge construction. Blended formats that integrate online resources with in-person sessions, along with student-centered frameworks such as technology-enabled active learning, have demonstrated improved outcomes by increasing interactivity, multimodal visualization, and structured inquiry. Within this pedagogical shift, virtual reality functions as an enabling medium that transforms learners from passive recipients into active participants by situating them in immersive, task-relevant contexts (Soliman et al., 2021). A growing body of research indicates that virtual reality can enhance cognitive and pedagogical outcomes in engineering domains, particularly where spatial reasoning and three-dimensional

10.3389/fcomp.2025.1701666

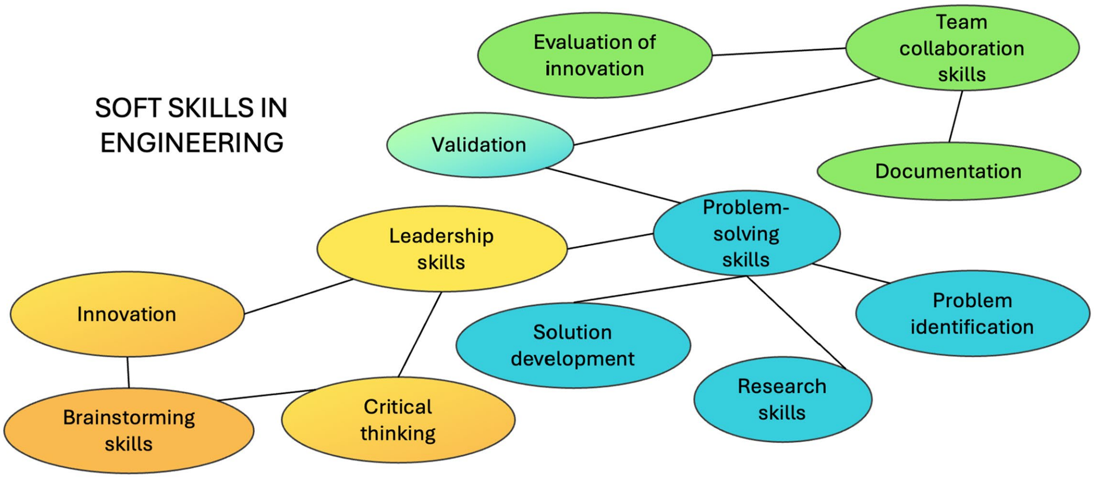

FIGURE 9
Soft skills in engineering.

phenomena are central. In courses such as fluid mechanics, virtual environments support comprehension of complex motion and flow patterns that are difficult to convey through two-dimensional diagrams or verbal description alone. Studies in electronics and three-dimensional visualization similarly report gains in achievement, engagement, critical thinking, and scientific attitudes (Soliman et al., 2021). There is also suggestive evidence that experiences encoded in virtual environments approximate real-world mnemonic processes more closely than conventional screen-based learning, potentially strengthening retention (Soliman et al., 2021).

On the other hand, the study proposes a teacher-support workflow that leverages a large language model to simulate teacher–student interactions from an initial lesson plan, derive reflective insights from those simulations, and then produce a revised, higher-quality plan that integrates both the process and the reflections. Using high-school mathematics as the testbed—Statistics, Functions, Algebra, and Geometry; the authors compiled 240 baseline plans (including LLM-generated plans with and without problem-chain prompts, expert-written plans, and plans from pre-service teachers) and created 240 "enhanced" plans by running one or two rounds of a simulation–reflection–refinement pipeline. Human raters evaluated all plans on an eight-point scale across nine categories and nineteen dimensions (Hu et al., 2025).

The enhanced plans consistently outperformed their baselines and often matched or exceeded expert plans. Gains were most pronounced in Statistics and weaker in Geometry, where text-only outputs struggled to convey diagram-intensive content. Introducing three problem-chain formats, context-based, trap-based, and summary-based, helped structure conceptual progression, although the more challenging trap-based questions remained a relative weakness. The approach was especially beneficial for pre-service teachers, improving practicality and alignment with lesson scope. Overall, the prompt-driven workflow lowers barriers to "pre-class rehearsal," supports human–AI co-design, and can be extended with multi-agent setups and human-in-the-loop review (Hu et al., 2025). These results suggest a promising opportunity to adapt the method to engineering laboratory courses as well.

Furthermore, Laboratory work remains fundamental to engineering education because it anchors theoretical constructs in inquiry and experimentation. However, physical laboratories are constrained by safety risks, infrastructure limitations, scheduling bottlenecks, and material costs. Virtual laboratories mitigate many of these constraints by providing safe, repeatable, and scalable practice without endangering learners or equipment. For distance learners, virtual laboratories can yield learning outcomes comparable to those achieved in traditional settings, thereby expanding access while maintaining educational quality. Accessibility can also improve for learners who benefit from seated operation and captioned content, although fully immersive head-mounted displays are not suitable for those with visual impairments. Cost remains a salient consideration: high-end systems provide greater immersion at a higher expense, whereas mobile solutions increase access at the cost of fidelity (Soliman et al., 2021).

Findings are not uniformly positive. Some studies report no statistically significant differences between virtual and traditional laboratories, while early implementations document student preferences for conventional laboratory experiences, citing perceived ease of operation and flexibility. Taken together, the literature supports a cautious yet optimistic conclusion: when virtual reality is thoughtfully integrated with lectures and online materials, it can enhance engineering education by making it more active and inquiry-driven, with particular advantages for topics involving complex spatial reasoning and for institutions seeking to broaden access to safe laboratory experiences. Effectiveness ultimately depends on careful instructional design, appropriate technological choices, and sensitivity to both learner and institutional contexts (Soliman et al., 2021).

Although this paper does not implement the proposed laboratory in a course, it describes how it could be implemented in engineering classes. In general, engineering laboratories lack a single definition or a standard program to follow (Hu et al., 2025). Since the advent of digital simulators, they have been used as instrumental systems to train engineers across diverse applications. Today, simulators can model highly complex systems with substantial fidelity, closely approximating real-world conditions. In this context, VR laboratories can integrate features of online laboratories and simulation laboratories to produce realistic environments that resemble a physical laboratory, complete with devices, a factory context, and expert avatars in specific domains.

The proposed VR laboratory would provide hands-on learning, enabling students to interact with equipment, domain experts in each section, and a virtual factory. The environment could be automatically adapted to different undergraduate levels, promoting dynamic interaction within the VR ecosystem. Rather than replicating a conventional, step-by-step laboratory where every student follows identical procedures, the design emphasizes iterative learning and formative feedback, eliminating the immediate pressure of summative grading. Flexibility and reliability would support access without fixed schedules or locations, as in remote-access laboratories (Steinemann and Braun, 2002). Because there is no rigid sequence of required exercises, the laboratory functions as a guided exploration space, allowing students to learn at their own pace. Moreover, VR implementation can reduce costs associated with physical maintenance and equipment failures. At the same time, the VR laboratory can incorporate both simulation-based exploration of theoretical concepts and experimental activities within the same environment, aiming to replicate the affordances of a real experimental laboratory (Steinemann and Braun, 2002; Balamuralithara and Woods, 2009).

Using the objectives of an engineering laboratory proposed by Feisel and Rosa (2005) as a reference, the fundamental objectives and main topics for the proposed virtual-reality laboratory using large language models are presented below. This course is a lab-centered pathway that encompasses the full lifecycle of a didactic extruder. Students can begin by applying first principles, decomposing subsystems, deriving governing relations, and mastering mechanics, electronics, power electronics, sensing, and multiloop control, while generating traceable data under explicit safety practices. They then utilize Expert Avatars, powered by large language models, to diagnose and improve the device, translate natural-language guidance into safe parameter and policy changes, conduct controlled experiments to isolate effects, perform failure analyses, and communicate results in an ethical manner. Finally, they apply these improvements in an advanced manufacturing context by preparing designs for manufacturability and assembly, integrating digital traceability, applying statistical process control, conducting controlled comparisons to support go/no-go decisions, and linking technical gains to takt time, capacity, cost per part, and return on investment. Across all phases, students utilize machine learning for maintenance, address energy and sustainability concerns, ensure accessibility, and maintain governance over changes and intellectual property, culminating in a factory-ready extruder improvement demonstrated through both technical and executive deliverables.

Phase 1: device-first foundations—the extruder and its components

Phase 1 begins by grounding students in the anatomy of the extruder and the first principles that govern it. Learners decompose the machine into its constituent parts: material feed, screw and barrel, heaters, drive train, die, and cooling. They then use this structure to derive relations for mass, momentum, energy, torque, throughput, and pressure. As they transition from plastic rheology and residence time to pressure–flow curves and torque–speed maps, the physics provides a practical lens for understanding overall efficiency. Those analytical bases carry directly into mechanics and structures, where students evaluate loads, tolerances, and fits for screws, barrels, bearings, couplings, and frames. By linking stress, strain, critical speed, thermal expansion, and mounting rigidity with observed noise and vibration, they learn to prevent wear, misalignment, and instability rather than reacting to it later.

With the mechanics stabilized, attention turns to thermal behavior. Students model heat flow across zones, design cooling paths, and tune profiles to deliver a stable melt with minimal overshoot. Concepts like conduction, convection, thermal lag, zoning strategy, insulation, and safety margin stop being abstract and become control levers for quality. The electrical layers then slot into the same systems picture. Learners read wiring diagrams, design safe low-voltage distribution and I/O systems, and apply grounding, shielding, and noise mitigation techniques to ensure that sensors and drives operate reliably. From there, they size and configure power electronics and drives, connecting rectification, heater control with pulse width, and variable-frequency or servo tuning to real current limits, harmonics, and protection schemes. This naturally leads to motion systems. Students select prime movers, gearboxes, and couplings by matching torque–speed envelopes to steady-state and transient demands, accounting for inertia, torque ripple, backlash, and alignment to ensure the chosen actuator can deliver what the process physics requires.

Sensing and control close the device loop. Learners select, place, calibrate, and condition sensors for temperature, pressure, torque, flow or throughput, and vibration, then synchronize and filter data for trustworthy analysis. Those signals feed multiloop control of temperature, motor speed, torque, and pressure, with interlocks and setpoint scheduling designed from the outset. The tuning work connects proportional–integral–derivative (PID) choices with feedforward, cascaded loops, anti-windup, and bumpless transfer, ensuring transitions are both responsive and safe. Throughout, safety and compliance are treated as design constraints, not afterthoughts. Students design guarding, emergency stops, interlocks, lockout and tagout awareness, safe states, thermal shielding, labeling, and markings so that risk reduction is documented and auditable. Phase 1 culminates in test and characterization. Teams plan structured experiments to measure throughput, energy use, melt quality, and stability, establish a golden batch, apply basic statistical process control, and produce clean, versioned datasets. By the end of the phase, device-level knowledge, control fluency, and data practice are woven into a single, testable whole.

Phase 2: expert avatars—language models in the loop

Phase 2 layers expert avatar guidance on top of the device foundations but keeps the human firmly in charge. Students first learn to design fast, targeted interactions that extract diagnostics, parameter suggestions, and design hypotheses without drifting into misleading responses. They use examples, bounded tool use, strict units, and versioned prompts so that advice remains traceable. The next step is translation. Avatar guidance is converted into safe, typed updates to parameters, policies, and design features using an application programming interface (API) schema with hard bounds, declared units, confidence reporting, rollback capabilities, and human approval. In other words, language becomes a precise control surface rather than an informal suggestion.

To ensure rigor, learners design avatar-in-the-loop experiments that separate the effect of advice from the effect of mechanical or control changes. They preregister hypotheses, randomize or counterbalance trials, plan manipulation checks, and confirm that sample sizes have adequate power. When failures occur, students diagnose them across sensors, mechanics, control logic, and avatar guidance, then update both the physical design and the avatar playbook. Fault trees, root-cause analysis, guardrails, post-mortems, and regression tests ensure that each fix prevents recurrence. Finally, communication and ethics make the process accountable. Results and decisions are reported with transparent instruction logs and explicit consent for the storage of transcripts. Bias checks are conducted on avatar outputs, and the findings are presented in clear executive summaries, accompanied by technical appendices, for both technical and non-technical audiences. By the end of Phase 2, learners have integrated trustworthy language-model collaboration into a safe engineering workflow.

Phase 3: factory integration, mass production, and machine learning

Phase 3 scales the improved device into a manufacturing reality. Students translate their designs into manufacturable variants using jigs and fixtures, assembly routings, and verified tolerances, thereby connecting engineering and manufacturing bills of materials with tolerance stacks, computer-aided manufacturing files, and printable fixtures, ensuring serviceability. These physical flows are mirrored in the digital factory. At the present stage, this digital factory has been implemented as a high-fidelity virtual model of the FrED line. Future work will expand it into a comprehensive digital twin by streaming real-time telemetry from the physical equipment and integrating in-situ control updates. Learners implement traceability by linking unit and lot genealogy to actual parts and changes, thereby maintaining as-built configurations in alignment with as-designed specifications and capturing approvals in change records. Quality systems then provide the statistical backbone. Teams could create control plans and apply statistical process control to critical characteristics, such as melt temperature, pressure stability, and dimensions. This involves selecting sampling plans and charts, computing capability indices, and handling nonconformances with discipline.

Validation at scale requires evidence, not anecdotes. Students conduct controlled comparisons between baseline and avatar-improved devices, ensuring adequate power, define uplift metrics and confidence intervals, and prepare lightweight submission packages for part approval when necessary. They connect technical improvement to production economics by translating changes into takt time, capacity, work in process, cost per part, and return on investment, using bottleneck analysis, quick changeover, sensitivity analysis, and make-or-buy decisions to quantify impact. Reliability and uptime are then addressed with machine learning. Learners engineer features from sensor data, build health indicators and remaining useful life estimates, set alert thresholds, and generate work orders, enabling maintenance to shift from reactive to predictive and prescriptive across heaters, bearings, screws, and drives. Energy and sustainability complete the operating picture. Students reduce energy consumption per unit and material waste while maintaining quality and takt, utilizing energy performance indicators, heater setpoint strategies, insulation improvements, compressed-air leak reduction, and improved material yield. The phase closes with handover and governance. Inclusive standard operating procedures and avatar-assisted training support adoption, while provenance and compliance are maintained through engineering change control, digital signatures and hashes, access control, audit trails, and regulatory mapping. What began as device-level mastery evolves into a factory-ready system, characterized by traceable decisions, measurable quality, and clear economics.

Creative Thinking Skills (CTS)—mapping and rubric for the VR–LLM laboratory

Additionally, this paper can be linked to the Creative Thinking Skills (CTS) criterion presented in Forte-Celaya et al. (2021), which provides concrete evidence generated by the VR–LLM laboratory and offers a ready-to-use rubric for scoring student work. Evidence is drawn from in-app artifacts (capstone plan, track deliverables), telemetry (logs of genetic algorithm and neural network), KPI charts, and the oral defense transcript.

How CTS maps to the laboratory

This laboratory advances Creative Thinking Skills (CTS) by guiding learners from exploratory optimization to justified, stakeholder-ready plans while capturing auditable evidence at every step. In the Line Production Optimizer, students tune buffer sizes and reliability policies using a genetic algorithm whose objective penalizes low throughput, oversized buffers, and off-baseline failure/repair rates. Because many parameter sets can still satisfy these constraints, there are multiple legitimate ways to succeed. Divergence, therefore, captures how far a learner's final operating plan departs from common or default settings while remaining viable, as shown by GA generation histories, the final policy vector, and the written rationale recorded in-app. That creative departure only matters if it benefits the factory, which leads directly to Impact: in the capstone, students defend their plan to avatar experts by balancing throughput, cost, stability, and maintainability. Clear KPI improvements, such as higher throughput, lower WIP, or better defect/energy profiles, combined with a persuasive oral defense, signal high Impact.

Crucially, novelty must fit the brief. The lab's three activity tracks and the Penta-S materials framework (smart, sustainable, social, sensing, safe) create space for uncommon yet well-justified solutions, such as pairing an ANN early warning system with a sustainability-driven buffer strategy. When a distinctive feature is introduced and explicitly tied to track objectives and Penta-S constraints, it earns Originality, evidenced in design notes, scenario choices, and justification screens. To make such novelty work in practice, learners must integrate ideas and tools; here, the lab's 14 interactive elements come into play. By chaining subsystem exploration, oscilloscope, and

CSV analysis, avatar consultations, GA optimization, and ANN forecasting into a single reasoning arc (signals → GA objective → ANN interpretation → plan), students demonstrate Flexibility, interdisciplinary synthesis visible in tool-use logs and the narrative rationale.

Quality of execution underpins all of the above. Because the lab exposes raw/processed signals, GA runs, and ANN outputs, raters can judge Technique directly: are objectives and penalties set correctly, are model outputs interpreted properly, and is telemetry used to validate decisions (avoiding pitfalls like overfitting or misreading failure/repair rates). Evidence appears in GA convergence curves, ANN accuracy/confusion matrices, correctly parameterized objectives, and clean data handling. Resolution then closes the loop by asking whether the learner actually met the brief defined in the FrED rubric—design quality, integration into manufacturing, performance and cost, innovation and problem-solving, and technology use—alongside track-level goals. Plans that meet throughput/cost targets, integrate subsystems coherently, and pass safety/feasibility checks earn high Resolution, as verified by threshold attainment, pass/fail checks, and subsystem integration snapshots.

The Creative Thinking Skills (CTS) Rubric for the VR–LLM Laboratory is presented in Table 2, based on Forte-Celaya et al. (2021).

Scoring: 1 = Beginning, 2 = Developing, 3 = Proficient, 4 = Exemplary. Evidence sources: in-app artifacts (capstone plan, track deliverables), GA/ANN/telemetry logs, KPI charts, oral defense transcript.

Total CTS score (max 24): sum of six criteria. Optional weighting (if emphasizing novelty/usefulness): Divergence, Impact, and Originality ×1.25; Flexibility, Technique, and Resolution ×1.0. Report both weighted and unweighted totals.

Results

The VR environment of the FrED Factory Learning Lab simulates a real-world factory setting. Figure 10 shows the spatial layout of the app, highlighting key training modules and activities.

The FrED Factory Learning Lab is available in two versions: for Android devices and for Oculus Meta Quest 3. While both versions aim to provide an engaging educational experience, the VR version supports full 3D-rendered avatars representing different engineering disciplines, whereas the Android version optimizes performance to prevent freezing by omitting avatars.

These platform-specific adjustments ensure accessibility and effectiveness for a diverse range of users. The VR version excels in immersion and interactivity, ideal for users with advanced hardware. The Android version offers portability, allowing engagement with training materials on mobile devices while maintaining core learning objectives.

Both versions focus on developing technical expertise and soft skills, offering tailored pathways for users to interact with FrED based on their device capabilities. This versatility makes the app suitable for a wide range of users, from on-the-go learners to those seeking a fully immersive experience.

TABLE 2 Creative Thinking Skills (CTS) rubric for the VR–LLM laboratory.

Criterion	4-exemplary	3-proficient	2-developing	1-beginning
Divergence	Reframes the operating policy with multiple defensible differentiators; departs clearly from common settings while satisfying all constraints (evidence: GA history + rationale).	At least one clear differentiator beyond typical solutions; plan remains viable.	Small/uncertain departure from defaults; novelty weakly justified.	Mirrors common/baseline settings; no meaningful differentiator.
Impact	Compelling KPI gains (throughput, cost, stability, maintainability) and a persuasive defense aligned to stakeholder needs; benefits are explicit and quantified.	Noticeable utility with trade-offs explained; most benefits quantified.	Some localized benefit; justification is partial or weakly quantified.	Limited or unclear value; benefits not demonstrated.
Originality	Distinctive feature(s) tightly aligned to track goals and Penta-S constraints; shows ideation breadth and careful selection.	At least one distinctive feature aligned with the brief and context.	Familiar ideas loosely tied to goals; limited justification.	No distinctive features; reproduces examples or defaults.
Flexibility	Integrates multiple tools/domains coherently (signals—GA objective—ANN inference—plan), with explicit handling of interfaces and trade-offs.	Purposeful combination of at least two tools/domains with clear intent.	Attempts cross-linking, but integration is weak or unclear.	Single-domain approach; no meaningful integration.
Technique	Rigorous setup and execution: correct objectives/penalties, valid model interpretation, sound data handling, reproducible runs (evidence: GA curves, ANN metrics).	Mostly correct methods with minor errors; interpretations generally sound.	Noticeable method or interpretation errors; partial mismatch to the task.	Frequent errors in setup/ interpretation; results not credible.
Resolution	Fully meets stated objectives/constraints (throughput/cost/safety/integration) with evidence; risks and limits acknowledged and managed.	Meets core objectives; minor gaps or untested constraints.	Partially meets objectives; key requirements are weak or missing.	Objectives largely unmet or unsupported by evidence.

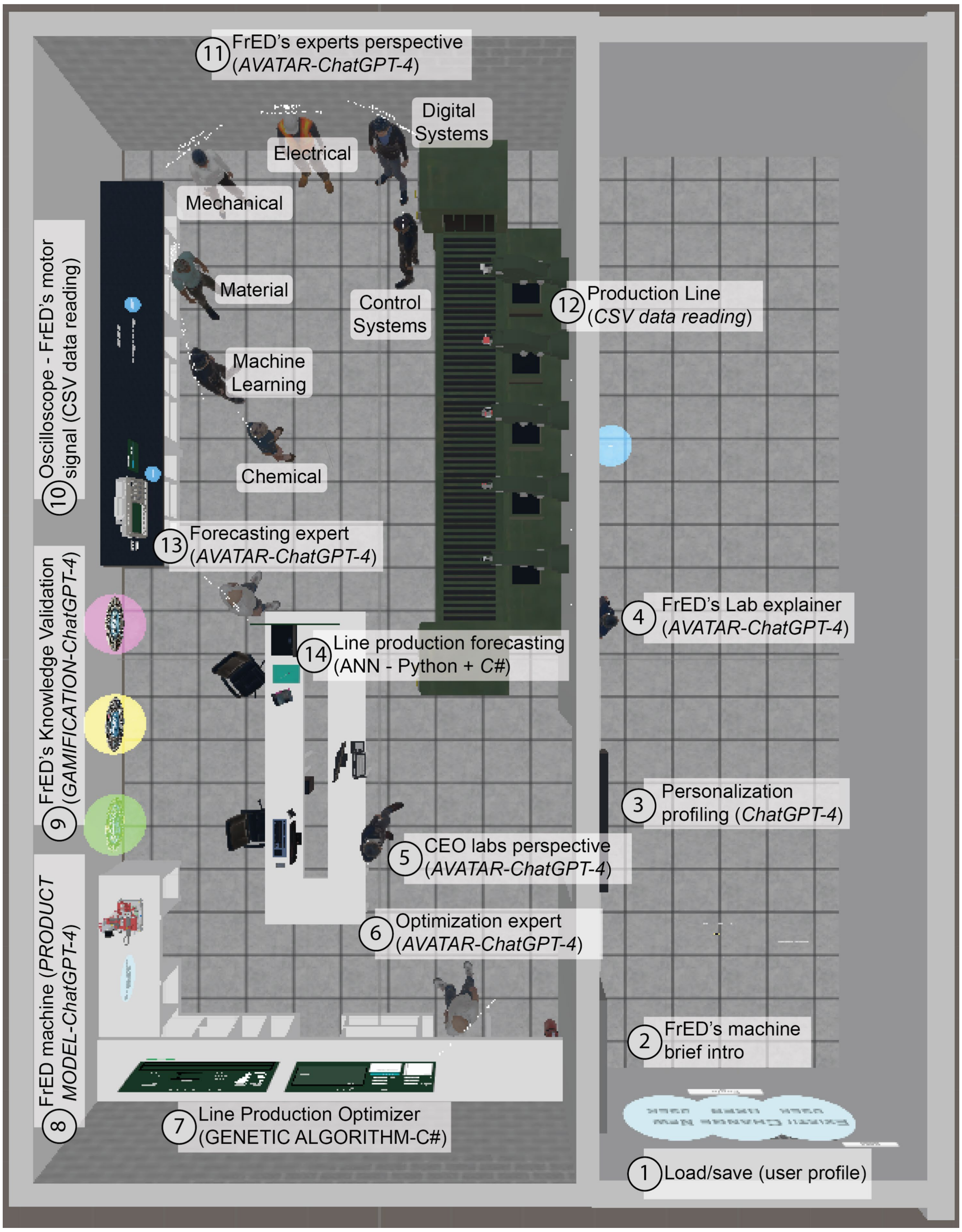

FIGURE 10
General layout of the proposed application.

Figure 11 provides snapshots of the Android version, showcasing platform accessibility, while Figure 12 presents the Oculus Meta Quest 3 interface, emphasizing the enhanced immersive experience of VR.

Users can interact with avatars or directly with FrED through voice commands or text input, catering to diverse communication preferences. Voice interaction offers a natural experience, while text input provides precision, making it suitable for quiet or noisy environments. ChatGPT also enables multilingual interactions, supporting both English and Spanish, enhancing inclusivity.

The platform features 14 interactive elements, each designed to meet specific training objectives and develop essential skills, including communication, leadership, problem-solving, and collaboration. Features range from user profile management to advanced modules, including predictive analytics utilizing genetic algorithms and neural networks.

1) Load/save user profile: allows users to save/load progress and profiles through local storage or Firebase, promoting organization and responsibility.

FIGURE 11
Different views of the 14 interactive features in the Android version.

2) FrED's machine brief intro: provides an overview of FrED's subsystems, fostering critical thinking by linking theory to machine operations.
3) Personalization profiling (ChatGPT-4): offers users a personality-based survey for tailored gamification or a default learning experience, fostering self-awareness and adaptability.
4) FrED's lab explainer (Avatar-ChatGPT-4): a virtual guide explains the lab environment and interactions, fostering communication and teamwork.
5) CEO labs perspective (Avatar-ChatGPT-4): offers a strategic view of lab operations, encouraging leadership and strategic thinking.
6) Optimizer expert (avatar-ChatGPT-4): guides users in applying optimization algorithms, enhancing problem-solving and innovative thinking.
7) Line production optimizer (Genetic Algorithm-C#): simulates production line optimization, promoting analytical thinking and research skills.
8) FrED machine (Product Model-ChatGPT-4): a detailed 3D model of FrED allows users to explore components, fostering critical thinking and problem-solving.
9) FrED's knowledge validation (Gamification-ChatGPT-4): challenges users with gamified activities to validate their knowledge, building confidence and adaptability.
10) Oscilloscope: FrED's motor signal (CSV data reading): provides real-time visualization of motor metrics, enhancing data analysis skills.
11) FrED's experts perspective (Avatars-ChatGPT-4): users interact with avatars from various disciplines, fostering interdisciplinary collaboration.
12) Production line (CSV data reading): simulates a production line for real-time data analysis, encouraging data-driven decision-making.
13) Forecasting expert (Avatar-ChatGPT-4): guides users on predictive analytics, enhancing forecasting and planning skills.

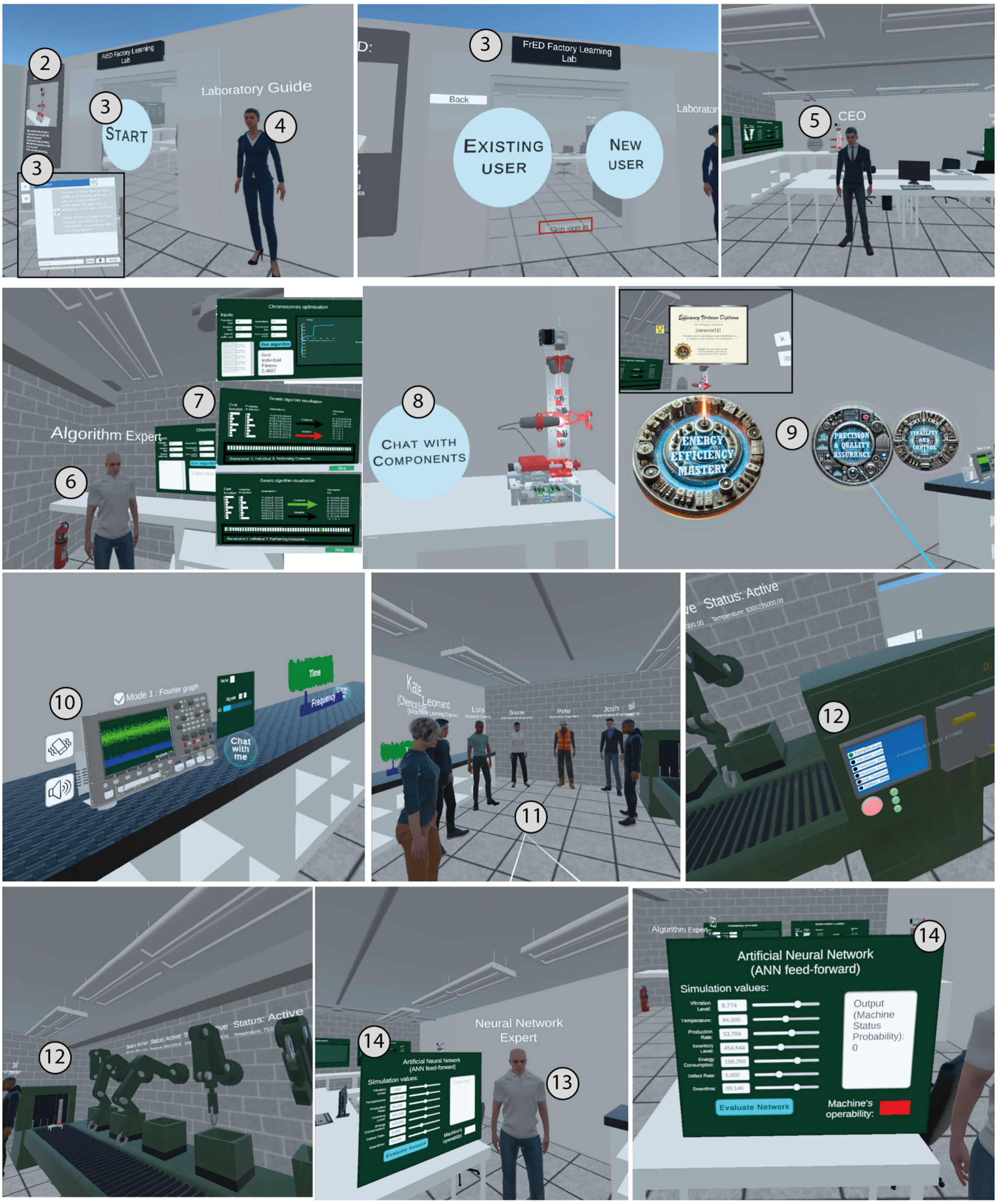

FIGURE 12
Different views of the 14 interactive features in VR.

14) Line production forecasting (ANN-Python+C#): trains users to predict operational statuses using an ANN, emphasizing analytical thinking and technical expertise.

Figure 13 summarizes the soft skills employed in each interactive feature.

The FrED Factory Learning Lab demonstrates how immersive virtual environments, gamification, and AI-driven personalization address evolving needs in engineering education. Leveraging both Android and VR platforms, the lab ensures inclusivity while maintaining core learning objectives, providing flexible and scalable solutions for users with varying hardware access.

The inclusion of 14 distinct features fosters critical thinking, problem-solving, collaboration, adaptability, and leadership—skills essential in modern engineering and manufacturing settings.

Feature	Soft skills
Load/save user profile	Organization, Responsibility, Self-management
FrED's machine brief intro	Critical Thinking, Problem Identification
Personalization profiling (ChatGPT-4)	Self-awareness, Decision-making, Adaptability
FrED's Lab explainer (AVATAR-ChatGPT-4)	Communication, Teamwork
CEO labs perspective (AVATAR-ChatGPT-4)	Leadership, Strategic Thinking, Innovation
Optimizer expert (AVATAR-ChatGPT-4)	Problem-solving, Analytical Thinking, Innovation
Line Production Optimizer (GENETIC ALGORITHM-C#)	Analytical Thinking, Research Skills, Decision-making
FrED machine (PRODUCT MODEL-ChatGPT-4)	Critical Thinking, Problem-solving, Collaboration
FrED's Knowledge Validation (GAMIFICATION-ChatGPT-4)	Confidence, Adaptability, Perseverance
Oscilloscope: FrED's motor signal (CSV data reading)	Attention to Detail, Data Analysis, Curiosity
FrED's experts perspective (AVATARS-ChatGPT-4)	Team Collaboration, Interdisciplinary Thinking
Production line (CSV data reading)	Critical Thinking, Problem-solving, Data-driven Decision-making
Forecasting expert (AVATAR-ChatGPT-4)	Forecasting, Planning, Innovation
Line production forecasting (ANN-Python+C#)	Analytical Thinking, Technical Expertise, Problem-solving

FIGURE 13
Soft skills are employed in each interactive feature.

Proposed plans and timeline for the empirical evaluation

Building on the framework outlined in Paszkiewicz et al. (2021), a systematic empirical evaluation shall be undertaken to assess its pedagogical effectiveness and usability in practical training environments. The assessment will concentrate on four primary dimensions:

Learning effectiveness

The first goal is to determine if the VR–LLM–gamified FrED laboratory delivers greater learning improvements than more traditional formats. Specifically, we will evaluate whether the system improves understanding of sustainable manufacturing, the FrED architecture and process flow, and the relationships between device-level settings (such as temperature, torque, reliability) and factory-level performance. These findings will be compared to results from standard instructional materials and a non-personalized VR version.

Soft-skills development

A second goal is to evaluate how effectively the laboratory supports the development of higher-order skills such as problem-solving, critical thinking, collaboration, and decision-making in production-line scenarios. Specifically, we will analyze how learners justify trade-offs within the Penta-S sustainability framework, design and defend optimized policies, and communicate their decisions to both technical and non-technical audiences.

User experience and acceptance

Third, we will examine learners' perceptions of immersion, usability, workload, and perceived usefulness, as well as their intention to adopt similar tools in the future. This involves analyzing how different system configurations (with or without LLM-based personalization and gamification) influence motivation, engagement, and perceived effectiveness.

Transfer and retention

Finally, we will examine whether the skills and knowledge acquired within the virtual laboratory are transferable to new scenarios, such as modified factory layouts or different failure patterns, and whether these competencies are retained over an extended period. This inquiry is particularly pertinent for industrial operators and engineering students, who are required to apply their learning to real or simulated production environments.

Study design

The empirical evaluation will be conducted using a controlled, between-subjects design. Participants will come from two primary populations: senior undergraduate and early graduate engineering students enrolled in courses related to manufacturing, control, or digital twins. Industrial operators or trainees working in sustainable manufacturing settings, where scenarios can be adapted to highlight safety, reliability, and process efficiency. To isolate the impact of VR, LLM-based guidance, and gamification, at least two experimental conditions will be implemented, with a third condition strongly recommended.

Treatment (T)

Full VR–LLM FrED laboratory, including personality-based profiles, adaptive avatars, gamified elements, and the three activity tracks (e.g., energy efficiency mastery, precision & quality assurance, stability & control).

Learners receive real-time, personalized guidance, and their actions are logged through the telemetry infrastructure already described (e.g., GA convergence, ANN performance, dialogue traces, tool usage).

Control 1 (C1)

Conventional instructional materials (slides, videos, and static simulations) covering the same core concepts, learning objectives, and approximate time-on-task as the treatment.

This condition represents a baseline close to current practice in many engineering programs.

Control 2 (C2) (recommended)

A non-personalized, non-gamified VR version of the laboratory, in which learners can explore and complete the same technical tasks without adaptive prompts or tailored game mechanics.

This condition allows us to isolate the added value of LLM-based personalization and gamification beyond the immersive VR medium itself.

Random assignment will be used wherever possible within each cohort (students and operators), and basic demographic and background information (e.g., prior VR experience, programming background, familiarity with FrED) will be collected to verify baseline comparability.

Measures and analysis

The evaluation will leverage both the assessment instruments and telemetry already designed for the virtual laboratory, as well as additional standardized questionnaires.

Knowledge and skills

Pre- and post-tests will assess understanding of FrED's system architecture and main subsystems.

Relationships between device-level parameters (e.g., extrusion temperature, torque, sensor placement, and controller tuning) and factory-level indicators (e.g., throughput, work-in-progress, reliability, and energy consumption).

Core concepts of the Penta-S sustainability framework and the role of optimization and forecasting in sustainable manufacturing.

Capstone performance will be evaluated using the existing FrED rubric and the CTS (critical thinking and technical skills) composite. Learners will solve a novel factory scenario and present a short oral defense of their decisions, allowing expert evaluators to rate decision quality, argumentation, and sustainability awareness.

Performance in VR tasks and system telemetry

Objective indicators will be extracted from system logs, such as:

- Task completion times and number of retries.
- Error rates, constraint violations, and time-to-stable policies in GA-based optimization.
- ANN forecasting accuracy and confusion matrices.
- Interaction with in-app tools (e.g., oscilloscope, CSV viewers) and avatar dialogue density and intent.
- These "mechanism traces" will help connect learning outcomes to how learners actually, explore the search space, interact with the avatars, and use the visualization tools.

User experience and acceptance

Standardized Likert-scale questionnaires will be administered after each intervention to measure:

- Perceived understanding (theoretical and practical).
- Motivation, engagement, and perceived effectiveness of each instructional form (traditional, non-personalized VR, VR–LLM–gamified).

Presence/immersion, usability, and workload (e.g., using instruments such as SUS, NASA-TLX, or presence scales).

Intention to use and perceived usefulness, drawing on technology acceptance models (e.g., TAM/UTAUT) in the operator study.

Soft skills and collaboration

Soft-skills rubrics will be developed to assess problem-solving, collaboration, and decision-making during the production-line scenarios (e.g., bottleneck resolution, buffer sizing, reliability trade-offs).

Expert raters will score team interactions and outcomes, focusing on how learners negotiate constraints, justify trade-offs, and communicate under time pressure.

Retention and transfer

A delayed post-test (4–6 weeks after the intervention) will be administered to a subset of participants to evaluate retention of key concepts and transfer to new scenarios, such as modified line configurations or different failure patterns.

Data analysis

Quantitative data (pre/post scores, telemetry metrics, questionnaire scales) will be analyzed using mixed ANOVA or linear mixed models to capture both within-subject changes and between-condition differences.

Group comparisons will test hypotheses about the added value of VR, LLM personalization, and gamification.

Qualitative data (open-ended responses and samples of LLM–avatar interactions) will be coded to identify common misconceptions, successful strategies, and user perceptions that numeric scales may not fully capture.

Timeline for 12-month implementation

The empirical evaluation is planned over approximately one year, in four main phases:

Months 1–2: design and preparation

During this phase, learning outcomes and soft skills indicators will be refined based on the curriculum and the existing FrED and CTS rubrics. Pre- and post-tests and questionnaires will be drafted, with items adapted from Paszkiewicz et al. (2021) on perceived effectiveness, recommendations, and understanding. Rubrics for collaboration and decision-making will be finalized. The experimental conditions (T, C1, optional C2), procedural scripts, and consent forms will be prepared, and the necessary ethical approvals confirmed.

Months 3–4: technical and pedagogical pilot

A small pilot study (approximately 10–15 students) will be conducted with one course group. The goal is to test VR logistics (headset deployment, network stability), task duration, user comfort, and instruction clarity. Initial questionnaires and short interviews will be collected, following a similar approach to previous VR deployments in education. Based on this feedback, task difficulty, user interface elements, and pacing will be refined, and equivalence of content and time-on-task across conditions will be checked.

Months 5–8: main controlled experiment with students

The primary student study will be carried out with approximately 30–40 participants per condition (subject to a power analysis and course enrollment). The intervention will be integrated into an existing course module so that the VR laboratory replaces or complements part of the traditional instruction. Pre-tests will be administered, followed by the assigned condition (T, C1, or C2). Post-tests, questionnaires, and telemetry data will then be collected. A delayed post-test will be scheduled 4–6 weeks later for retention and transfer analysis.

Months 9–10: extension to operators and external cohorts

In this phase, the scenarios and instruments will be adapted for industrial operators, with a focus on safety, reliability, and operational decision-making. A smaller but ecologically rich study (around 15–25 participants) will be run, typically comparing the treatment condition to one of the control conditions. User-acceptance measures (e.g., intention to use, perceived usefulness) will be included. Comparisons will be made between student and operator populations to explore generalizability and differences in usage patterns.

Months 11–12: analysis and iteration

The final phase will focus on comprehensive statistical analysis and triangulation of quantitative results with qualitative feedback and interaction logs. Based on the findings, the FrED VR laboratory will be refined (e.g., adjusting scenario complexity, feedback mechanisms, and gamification loops). The goal is also to package the evaluation framework so it can be reused in future multi-site studies across different campuses and institutions.

Discussion

The combination of VR, large language models, and AI has significant potential for transforming education and training in sustainable manufacturing. VR provides hands-on experiences, helping learners understand complex manufacturing processes without the need for physical machinery, thereby reducing costs and risks compared to traditional training methods. This aligns with Industry 4.0 and 5.0, which focus on enhancing efficiency while prioritizing worker empowerment and education.

By leveraging genetic algorithms and neural networks, the VR lab simulates production scenarios and optimizes key parameters, enabling learners to visualize the impact of various factors in real-time. The iterative optimization process helps users understand the balance required in industrial settings, including production rates, machine reliability, and cost efficiency.

The VR lab emphasizes the Penta-S approach, contributing to understanding sustainable practices by demonstrating efficient organization and resource management. This focus on sustainability is timely given the global push to reduce environmental impact in manufacturing.

Despite the advanced capabilities of VR and AI, challenges such as accessibility and the learning curve must be addressed to maximize their benefits. Ensuring a user-friendly VR environment and providing sufficient training are key to overcoming these challenges.

The use of a loss function that penalizes inefficiencies highlights the emphasis on efficiency but may limit flexible, creative exploration that learners sometimes need. Expanding the VR lab to include real-time collaboration among users in different locations could make it a valuable global educational resource.

The adaptability of large language models to provide real-time guidance in the VR lab enhances the learning experience through personalized support, using text or voice interfaces to cater to diverse learning preferences.

In designing the FrED VR–LLM lab, the user interface, scaffolding mechanisms, and pacing are intentionally tailored to boost presence while minimizing cognitive overload. First, the UI follows cognitive load theory and CTML principles by removing unnecessary, purely decorative elements and showing only task-relevant information, since excessive visual stimulation and ornamental details in VR can hinder learning despite increasing presence (Sari et al., 2024). In practice, this is achieved through a small number of stable, color-coded panels (device view, parameter sliders, oscilloscope/CSV viewers, and GA/ANN dashboards) that are reused across the three core modules—FrED device, avatar guidance, and factory—so learners do not have to relearn the interface when moving from the extruder to the production line.

A simplified, performance-optimized Android version with fundamental interactions (no 3D avatars, same tasks and metrics) is available for lower-end devices or beginners, aligning with evidence that accessible, easy-to-use interfaces and straightforward menus help reduce cognitive load and improve perceived learning effectiveness in immersive VR.

Instructional scaffolding is provided through both the activity structure and the LLM-driven avatars. The training program is divided into three VR activities (Energy Efficiency Mastery, Precision and Quality Assurance, Stability and Control), each focusing on a specific subsystem of FrED and increasing in complexity.

Within each activity, learners receive step-by-step prompts, immediate feedback, and gamified scoring with sealed diplomas that show progress without overwhelming the interface with raw data.

Avatars like the Lab Explainer, CEO perspective, Optimizer, and Forecasting experts serve as just-in-time tutors: they translate domain jargon into natural language, suggest bounded parameter edits, and prompt reflection on trade-offs between throughput, reliability, and sustainability, rather than just adding more information. This guided experimentation approach echoes research on self-presence and avatar design, where carefully tuned embodiment and feedback (e.g., synchronized movements, consistent responses) support cognitive performance without unnecessary sensory complexity (Jahn et al., 2019).

To explicitly reduce cognitive overload, the FrED lab adopts three complementary strategies supported by recent empirical work on VR and cognitive load. First, task–technology fit is ensured by making sure every interactive feature (such as line optimizer, oscilloscope, knowledge-validation mini-games) directly supports a specific learning goal, avoiding tempting but irrelevant details that would increase extraneous load. Second, complexity is increased gradually across phases, from device-level tests to avatar-guided diagnostics to factory-level GA/ANN optimization, so the intrinsic load grows slowly, and learners can reuse earlier schemas instead of starting from scratch in the factory module. Third, during pilots, interaction logs and questionnaires inspired by cognitive overload studies (e.g., ratings of task difficulty and mental effort) are used to identify screens or sequences that produce high load. These are then redesigned with simpler menus, fewer simultaneous elements, or additional avatar prompts. All these UI and pedagogical choices aim to keep FrED's immersive experience in the utilitarian zone, supporting reflection and skill transfer, rather than letting presence and stimulation undermine learning effectiveness.

Future work

Although the proposed VR laboratory integrates advanced technologies such as LLMs, gamification, and optimization algorithms, it has not yet been validated through a systematic user study. Future work will include a controlled experiment to evaluate the educational effectiveness of the system. A planned design involves a between-subjects study with a treatment group (VR–LLM laboratory with personalization and gamification) and a control group (standard materials or a non-personalized VR version). Participants will include senior undergraduate and early graduate engineering students, as well as operators, ensuring comparable baseline preparation.

Conclusion

This proposal highlights the development of a VR laboratory powered by large language models and AI, demonstrating their role in enhancing sustainable manufacturing through the FrED system. FrED serves as an AI-driven framework for discovering and optimizing Penta-S materials, products, and sustainable practices. A key contribution of this work is the application of AI to optimize production parameters, including buffer size, production rate, and machine reliability, as well as the use of neural networks for forecasting. This approach maximizes efficiency while minimizing costs and downtimes.

The optimization process results are presented through visual feedback, which shows the evolution of key parameters across generations, providing insight into the progression toward more efficient solutions. A similar approach is used for neural network-based forecasting.

This paper also emphasizes the educational applications of VR and AI, with the potential to transform learning in sustainable engineering. By integrating VR and AI-assisted education, the training process becomes more interactive and engaging, helping industry professionals and engineering students master sustainable manufacturing practices. Overall, this paper demonstrates the potential of AI and genetic algorithms in advancing sustainable manufacturing, while highlighting how VR and AI technologies can enhance education and training.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Gabriela Torres Delgado, Secretary of the Institutional

Committee on Research Ethics, Tecnologico de Monterrey. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PP: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. BA: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. JM: Conceptualization, Data curation, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing, Methodology, Visualization. RB: Conceptualization, Data curation, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing, Resources, Supervision. JG: Conceptualization, Data curation, Investigation, Software, Writing – original draft, Writing – review & editing. OM: Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported in part by the MIT–Tecnológico de Monterrey Program in Nanoscience and Nanotechnology. Also, it was supported by Institute of Advanced Materials for Sustainable Manufacturing, Tecnologico de Monterrey.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2025.1701666/full#supplementary-material

References

Abdelouahab, A. (2020). Virtual reality–enhanced soft and hard skills development environment for higher education. In *The Proceedings of the Third International Conference on Smart City Applications* (pp. 255–267). Cham: Springer International Publishing.

Acevedo, P., Magana, A. J., Benes, B., and Mousas, C. (2024). A systematic review of immersive virtual reality in STEM education: advantages and disadvantages on learning and user experience. *IEEE Access* 12, 1–28. doi: 10.1109/ACCESS.2024.3489233

Alsaleh, S., Tepljakov, A., Köse, A., Belikov, J., and Petlenkov, E. (2022). Reimagine lab: bridging the gap between hands-on, virtual and remote control engineering laboratories using digital twins and extended reality. *IEEE Access* 10, 89924–89943. doi: 10.1109/ACCESS.2022.3199371

Balamuralithara, B., and Woods, P. C. (2009). Virtual laboratories in engineering education: the simulation lab and remote lab. *Comput. Appl. Eng. Educ.* 17, 108–118. doi: 10.1002/cae.20186

Bezanson, K., Soberanis, L., Thomas, B., Brooks, R., and Rojas-Muñoz, E. (2023). Towards an intelligent tutoring system for virtual reality learning environments. In *2023 IEEE Frontiers in Education Conference (FIE)*. Texas, USA: IEEE. 1–5.

Botke, J. A., Jansen, P. G., Khapova, S. N., and Tims, M. (2018). Work factors influencing the transfer stages of soft skills training: a literature review. *Educ. Res. Rev.* 24, 130–147. doi: 10.1016/j.edurev.2018.04.001

Bradley, R. (2023). Design and manufacturing of educational fiber extrusion device and smart factory. Cambridge, MA: Massachusetts Institute of Technology, Department of Mechanical Engineering.

Caeiro-Rodríguez, M., Manso-Vázquez, M., Mikic-Fonte, F. A., Llamas-Nistal, M., Fernández-Iglesias, M. J., Tsalapatas, H., et al. (2021). Teaching soft skills in engineering education: an European perspective. *IEEE Access* 9, 29222–29242. doi: 10.1109/ACCESS.2021.3059516

Chen, L., Cai, Y., Wang, R., Ding, S., Tang, Y., Hansen, P., et al. (2024). Supporting text entry in virtual reality with large language models. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, Orlando, Florida, USA: IEEE. 524–534.

Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: a review. *IEEE Access* 8, 75264–75278. doi: 10.1109/ACCESS.2020.2988510

de Giorgio, A., Gargiulo, F., Greco, A., Reina, G., and D'Addona, D. M. (2023). Adopting extended reality? A systematic review of opportunities, challenges and practices in manufacturing education. *J. Vocat. Behav.* 146:103858. doi: 10.1016/j.jvb.2023.103858

El Hajji, M., Ait Baha, T., Berka, A., Ait Nacer, H., El Aouifi, H., and Es-Saady, Y. (2025). An architecture for intelligent tutoring in virtual reality: integrating LLMs and multimodal interaction for immersive learning. *Information* 16:556. doi: 10.3390/info16070556

Fazlollahi, A. M., Bakhaidar, M., Alsayegh, A., Yilmaz, R., Winkler-Schwartz, A., Mirchi, N., et al. (2022). Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. *JAMA Netw. Open* 5:e2149008. doi: 10.1001/jamanetworkopen.2021.49008

Feisel, L. D., and Rosa, A. J. (2005). The role of the laboratory in undergraduate engineering education. *J. Eng. Educ.* 94, 121–130. doi: 10.1002/j.2168-9830.2005.tb00833.x

Filho, P. T. J. C., and Junior, O. C. (2025). Digital Twin in Industry 4.0: Systematic review and content analysis and an architectural reference model. *Proce. Com. Sci.* 253, 2919–2928. doi: 10.1016/j.procs.2025.02.016

Flores Romero, P., Shaw, R. S., Hirsh, D., and Pirolli, P. (2025). Structured human-LLM interaction design reveals information foraging patterns in pedagogical content creation. *Humanit. Soc. Sci. Commun.* 12:405. doi: 10.1057/s41539-025-00332-3

Forte-Celaya, J., Ibarra, L., and Glasserman-Morales, L. D. (2021). Analysis of creative thinking skills development under active learning strategies. *Educ. Sci.* 11:621. doi: 10.3390/educsci11100621

Frankford, E., Sauerwein, C., Bassner, P., Krusche, S., and Breu, R. (2024). AI-tutoring in software engineering education. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training* (pp. 309–319).

Hamid, N. S. S., Aziz, F. A., and Azizi, A. (2014). Virtual reality applications in manufacturing system. In *2014 Science and Information Conference*, London, United Kingdom: IEEE. 1034–1037.

Hassan, M., Montague, G., Iqbal, M. Z., and Fahey, J. (2024). Virtual reality-based bioreactor digital twin for operator training. *Digit. Chem. Eng.* 5:100101. doi: 10.1016/j.dche.2024.100101

Hickman, L., and Akdere, M. (2017). "Exploring virtual reality for developing soft-skills in STEM education" in 2017 7th world engineering education forum (WEEF). Kuala Lumpur, Malaysia: IEEE, 461–465.

Hoang, T., Greuter, S., and Taylor, S. (2022). An evaluation of virtual reality maintenance training for industrial hydraulic machines. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Christchurch, New Zealand: IEEE, 573–581.

Hu, B., Zhu, J., Pei, Y., and Gu, X. (2025). Exploring the potential of LLM to enhance teaching plans through teaching simulation. *NPJ Sci. Learn.* 10:7. doi: 10.1038/s41539-025-00300-x

Ipsita, A., Kaki, R., Liu, Z., Patel, M., Duan, R., Deshpande, L., et al. (2025). Virtual reality in manufacturing education: a scoping review indicating state-of-the-art, benefits, and challenges across domains, levels, and entities. arXiv Available online at: https://arxiv.org/abs/2503.18805

Jahn, K., Kordyaka, B., Ressing, C., Roeding, K., and Niehaves, B. (2019). "Designing self-presence in immersive virtual reality to improve cognitive performance—a research proposal" in Information systems and neuroscience: NeuroIS retreat 2019 (Cham: Springer International Publishing), 83–91.

Junior, O. C., Moraes, R., and dos Santos, A. (2025). Digital twin in industry 4.0: systematic review and content analysis. *Procedia Comput. Sci.* 242, 1667–1677. doi: 10.1016/j.procs.2025.03.207

Karanam, S. A. K., and Hartman, N. W. (2025). A systematic review of digital twin and virtual learning environments for smart manufacturing education. *Manuf. Lett.* 44, 1597–1608. doi: 10.1016/j.mfglet.2025.06.179

Khelifi, S., and Morris, A. (2024). Mixed reality IoT smart environments with large language model agents. In *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*. Toronto Canada: IEEE, 1–7.

Kim, W. H., and Kim, J. H. (2020). Individualized AI tutor based on developmental learning networks. *IEEE Access* 8, 27927–27937. doi: 10.1109/ACCESS.2020.2972167

Kumar, V. V., Carberry, D., Beenfeldt, C., Andersson, M. P., Mansouri, S. S., and Gallucci, F. (2021). Virtual reality in chemical and biochemical engineering education and training. *Educ. Chem. Eng.* 36, 143–153. doi: 10.1016/j.ece.2021.05.004

Lampropoulos, G., Fernández-Arias, P., de Bosque, A., and Vergara, D. (2025). Virtual reality in engineering education: a scoping review. *Educ. Sci.* 15:1027. doi: 10.3390/educsci15081027

Létourneau, A., Deschênes, R., and Lajoie, S. P. (2025). A systematic review of AI-driven intelligent tutoring systems in K-12 education: effects and evaluation designs. *Educ. Inf. Technol.* doi: 10.1007/s10639-025-13269-0

Li, S. (2022). "Design of Intelligent Manufacturing Training Platform Combining Virtual and Real" in 2022 3rd Asia symposium on signal processing (ASSP) (Singapore: IEEE), 92–95.

Liu, Z., Agrawal, P., Singhal, S., Madaan, V., Kumar, M., and Verma, P. K. (2025). LPITutor: an LLM based personalized intelligent tutoring system using RAG and prompt engineering. *PeerJ Comput. Sci.* 11:e2991. doi: 10.7717/peerj-cs.2991

Liu, Z., Chen, D., Zhang, C., and Yao, J. (2024). Design of a virtual reality serious game for experiencing the colors of Dunhuang frescoes. *Herit. Sci.* 12:370. doi: 10.1186/s40494-024-01477-x

Molina, A., Cortés, D., Chairez, I., Alfaro-Ponce, M., Alvarez, M. M., and Trujillo de Santiago, G. (2024). Comprehensive sustainable development of a multifunctional machine: 3D food printer and didactic platform. *Int. J. Sustain. Eng.* 17, 413–428. doi: 10.1080/19397038.2024.2355895

Molina, A., Zúñiga, A. E., Medina, D. I., Zaldívar, F. G., Méndez, J. I., Ponce, P., et al. (2025). Digital technologies for sustainable manufacturing. Boca Raton, USA: CRC Press.

Oliveira, D. M., Cao, S. C., Hermida, X. F., and Rodríguez, F. M. (2007). "Virtual reality system for industrial training" in 2007 IEEE international symposium on industrial electronics. (Vigo, Spain: IEEE), 1715–1720.

Oliveira, W., Hamari, J., Shi, L., Toda, A. M., Rodrigues, L., Palomino, P. T., et al. (2023). Tailored gamification in education: a literature review and future agenda. *Educ. Inf. Technol.* 28, 373–406. doi: 10.1007/s10639-022-11122-4

Paszkiewicz, A., Salach, M., Strzałka, D., Budzik, G., Nikodem, A., Wójcik, H., et al. (2021). VR education support system—a case study of digital circuits design. *Energies* 15:277. doi: 10.3390/en15010277

Peniche, A., Diaz, C., Trefftz, H., and Paramo, G. 2012). Combining virtual and augmented reality to improve the mechanical assembly training process in manufacturing. In *Proceedings of the 2012 American conference on applied mathematics* (pp. 292–297)

Pester, A., Tammaa, A., Gütl, C., Steinmaurer, A., and Abou El-Seoud, S. (2024). Conversational agents, virtual worlds, and beyond: a review of large language models enabling immersive learning. In *2024 IEEE Global Engineering Education Conference (EDUCON)*. Kos, Greece: IEEE, 1–6.

Peterson, L., Gosea, I. V., Benner, P., and Sundmacher, K. (2025). Digital twins in process engineering: an overview on computational and numerical methods. *Comput. Chem. Eng.* 193:108917. doi: 10.1016/j.compchemeng.2024.108917

Pirola, C., Peretti, C., and Galli, F. (2020). Immersive virtual crude distillation unit learning experience: the EYE4EDU project. *Comput. Chem. Eng.* 140:106973. doi: 10.1016/j.compchemeng.2020.106973

Ponce, H., and Ponce, P. 2011). Artificial organic networks. In *2011 IEEE Electronics, Robotics and Automotive Mechanics Conference*. Cuernavaca, Morelos, Mexico: IEEE, 29–34.

Ponce-Cruz, P., Gutiérrez, A. M., Ramírez-Mendoza, R. A., Flores, E. M., Espinoza, A. A. O., and Silva, D. C. B. (2020). A practical approach to metaheuristics using LabVIEW and MATLAB®. New York, USA: Chapman and Hall/CRC.

Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput. Educ.* 147:103778. doi: 10.1016/j.compedu.2019.103778

Rosko, R. S. 2024 "FrED Manufacturing - A Study in Affordable Manufacturing to Scale using Desktop Sized Fiber Extrusion Device." Available online at: https://dspace.mit.edu/handle/1721.1/157212 (Accessed 11, 2024).

Rubio, E. M., Sanz, A., and Sebastián, M. A. (2005). Virtual reality applications for the next-generation manufacturing. *Int. J. Comput. Integr. Manuf.* 18, 601–609. doi: 10.1080/09511920500069259

Sari, R. C., Pranesti, A., Solikhatun, I., Nurbaiti, N., and Yuniarti, N. (2024). Cognitive overload in immersive virtual reality in education: more presence but less learnt? *Educ. Inf. Technol.* 29, 12887–12909. doi: 10.1007/s10639-023-12379-z

Schofield, D. (2012). Mass effect: a chemical engineering education application of virtual reality simulator technology. *J. Online Learn. Teach.* 8:63.

Shoa, A., and Friedman, D. (2025). Milo: an LLM-based virtual human open-source platform for extended reality. *Front. Virtual Real.* 6:1555173. doi: 10.3389/frvir.2025.1555173

Soliman, M., Pesyridis, A., Dalaymani-Zad, D., Gronfula, M., and Kourmpetis, M. (2021). The application of virtual reality in engineering education. *Appl. Sci.* 11:2879. doi: 10.3390/app11062879

Steinemann, M. A., and Braun, T. (2002). Remote versus traditional learning in a computer networks laboratory, communications and computer networks. Cambridge, MA, 503–507.

Xu, W. (2024). Development and production of affordable desktop fiber extrusion devices (fred) for educational purposes. Massachusetts, USA: Massachusetts Institute of Technology, Department of Mechanical Engineering.

Yang, Y., Deb, S., He, M., and Kobir, M. H. (2023). The use of virtual reality in manufacturing education: state-of-the-art and future directions. *Manuf. Lett.* 35, 1214–1221. doi: 10.1016/j.mfglet.2023.07.023

Yang, C., Zhang, J., Hu, Y., Yang, X., Chen, M., Shan, M., et al. (2024). The impact of virtual reality on practical skills for students in science and engineering education: a meta-analysis. *Int. J. STEM Educ.* 11:28. doi: 10.1186/s40594-024-00487-2

Zhang, Y. 2024 "FrED Manufacturing—A Study in Affordable Manufacturing to Scale Using Desktop Sized Fiber Extrusion Device." Available online at: https://dspace.mit.edu/handle/1721.1/157159?show=full