



Embedding stealth assessment in metaverse-based training for PC building: advancing technical education using evidence-centered design

Manuel B. Garcia^{1,2,3} 

Received: 24 January 2025 / Accepted: 17 June 2026

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract

Technical and Vocational Education and Training (TVET) programs require reliable methods to teach and assess performance-based skills in authentic contexts. Immersive platforms such as the metaverse offer promising environments for simulating real-world tasks and providing scalable training opportunities. In these technology-rich settings, stealth assessment is recognized as a psychometrically sound and theoretically grounded approach. However, its application has been largely confined to video games and primarily in contexts emphasizing creativity and collaboration. This study broadens that scope by embedding stealth assessment within a metaverse environment to develop and evaluate technical competencies in PC building. Guided by evidence-centered design, a two-part investigation examined (1) the suitability of stealth assessment in a metaverse context and (2) the effectiveness of metaverse-based training in improving learning outcomes and facilitating skill transfer. Undergraduate students from computing and engineering programs engaged in structured PC-building scenarios within the metaverse, where their performance was tracked through automated event logs, validated against expert evaluations, and analyzed using inferential statistical methods. Results indicate that embedding stealth assessment in the metaverse enabled reliable and unobtrusive evaluation of learners' technical competencies. Metaverse-based PC-building training yielded higher learning gains than traditional methods, and learners demonstrated the ability to transfer skills from virtual tasks to physical PC-building. Overall, these findings offer evidence-informed insights for educators, researchers, and policymakers on integrating immersive technologies and embedded assessments in technical education and vocational training.

Keywords TVET · Technical education · PC Building · Evidence-centered design · Stealth assessment · Metaverse · MILES virtual world

Extended author information available on the last page of the article

Published online: 03 July 2026

 Springer

1 Introduction

Technical education is a cornerstone of modern life and work in a technology-powered world. As the global workforce increasingly relies on technological tools and digital platforms, understanding and working with technology has become a basic competency across industries (Beer & Mulder, 2020; Li, 2024). A systematic literature review highlights technical skills as one of three essential employability skill types, along with cognitive and non-cognitive skills, critical for Industry 4.0 and beyond (Fajaryati et al., 2020). Supporting this viewpoint is the recent report from the World Economic Forum (2025), which underscores that technical proficiency remains integral to the core skills demanded across industries and regions globally. While big data and artificial intelligence (AI) top the list of the fastest-growing skills, there remains a critical need to reinforce foundational technical competencies. One example of a foundational competency is the capability to assemble, maintain, optimize, and troubleshoot personal computers (PCs)—henceforth referred to as PC building. Among other technical know-how, possessing PC building expertise means that individuals can identify suitable components, ensure system compatibility, resolve technical issues, and ensure optimal performance for specific purposes (e.g., gaming, productivity, or general use). The collective ability to perform these tasks reflects underlying competencies such as critical thinking, problem-solving, decision-making, technical knowledge, attention to detail, and adaptability. Mastering these skills prepares individuals for the dynamic and unpredictable nature of technological advancements (Poláková et al., 2023; Tushar & Sooraksa, 2023). With PC building as a gateway to mastering advanced technologies and securing opportunities in a competitive labor market, there is a necessity to develop and implement effective strategies for teaching and assessment in technical education.

Within formal education and workforce development, PC building is best situated within the broader area of computer hardware servicing (Bringula et al., 2022). This field encompasses the installation, configuration, maintenance, and troubleshooting of computer systems and networks, positioning it as a cornerstone of Information and Communications Technology (ICT) programs. As a formally recognized qualification area, this set of competencies is embedded in Technical and Vocational Education and Training (TVET) frameworks that align curricula with occupational standards, certification requirements, and labor market demands. International guidelines such as *TVET Strategy 2022–2029* (UNESCO, 2022) and the *Global Framework on Core Skills for Life and Work* (International Labour Organization, 2021) highlight the need for digital competence, hardware literacy, and problem-solving as essential for workforce readiness in digital economies. Recent TVET literature recommends competency-based whole-institution approaches to digital competence to ensure leaders, educators, and learners all share responsibility for technical skills development (e.g., Zhong & Juwaheer, 2024). In this environment, PC building exemplifies the foundational technical proficiency that supports progression toward advanced ICT qualifications and certification. Framing PC building in this way consequently highlights its dual role as a foundational technical practice and as an integral component of structured pathways toward employability, labor mobility, and lifelong learning in technology-driven economies.

1.1 Stealth assessment and evidence-centered design

Stealth assessment is a covert method of evaluating learner competencies by embedding assessment processes seamlessly into interactive environments. In contrast to traditional assessments, which often disrupt learning and induce test anxiety (Jerrim, 2023), stealth assessment operates unobtrusively by capturing performance data as learners engage with tasks in real time (Shute et al., 2009). This approach allows for continuous monitoring of competency levels and supports adaptive interventions such as adjusting task difficulty or delivering personalized feedback. Importantly, stealth assessment should not be understood as a form of deception but as a means of invisibly collecting gameplay data and applying it formatively to support performance gains, foster self-regulation, and nurture independent learning. Such an approach directly tackles the problem articulated by Shute et al. (2016b), particularly on capturing evidence of learning unobtrusively and transforming it into actionable feedback without disrupting engagement. Consequently, several studies have deployed and evaluated the use of stealth assessment in video game contexts (e.g., Fang et al., 2023; Shute & Rahimi, 2021). Despite these efforts, a systematic literature review on the use of video games noted that the development and application of stealth assessment remain in a relatively dormant state (Varghese & Renumol, 2024). This gap suggests the need for further inquiry and broader implementation to unlock the full potential of stealth assessment as a mechanism for evaluating competencies.

A common conceptual framework that provides theoretical foundation for stealth assessment is Evidence-Centered Design (ECD). ECD is a systematic approach to assessment design that focuses on aligning the tasks, observations, and interpretations used in assessments to measure specific competencies or skills (Mislevy et al., 2003). It provides a framework for ensuring that assessments are both valid and reliable by explicitly defining what is being measured, how it is measured, and why the measurements are meaningful (Newton et al., 2021). Although originally developed for educational and psychological testing, ECD has since been applied across diverse domains. At its core, ECD is structured around three interrelated models (see Fig. 1). The competency model (or student model) specifies the knowledge, skills, and attributes targeted for assessment. In the context of computer programming education, for instance, this model may include theoretical understanding, algorithmic thinking, logical reasoning, and other proficiencies essential for programming expertise (Garcia, 2024). The evidence model then identifies the behaviors and performances that

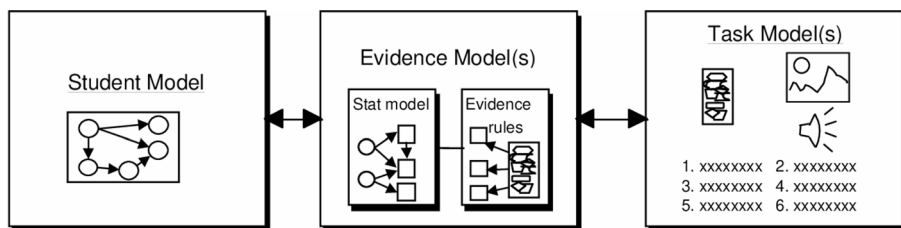


Figure 1 The Central Models of the Conceptual Assessment Framework

Fig. 1 Three Interrelated Models of ECD (Mislevy, 2013)

serve as indicators of those competencies. Such evidence may be drawn from learners' interactions within coding environments, including error frequency, time taken to solve problems, or the creativity of implemented solutions. Finally, the task model prescribes the design and structure of assessment activities that elicit the required evidence. In programming education, this model might include writing a sorting algorithm, debugging existing code, or completing a project with specified functional requirements. Collectively, these models establish a coherent design logic that not only strengthens the interpretive validity of assessment results but also provides a robust theoretical foundation for embedding stealth assessment into complex learning environments.

1.2 Video games, digital environments, and metaverses

According to Rahimi and Shute (2024), stealth assessment is a theoretically-grounded and psychometrically-reliable approach for evaluating, enhancing, and exploring learning processes within technology-enhanced environments. Among interactive platforms, video games (e.g., Fang et al., 2023) and game-based learning environments (e.g., Min et al., 2020) stand out as particularly popular mediums for implementing stealth assessment. A shared characteristic of these mediums is their inherent enjoyment (Caroux & Pujol, 2024) and immersive experiences (Christou, 2014) that foster high levels of engagement and motivation. Such characteristics are critical in learning contexts, as students who are motivated and immersed in the process are more likely to stay focused, persist through challenges, and deeply engage with the content (Bowden et al., 2021; Clark et al., 2016; Li et al., 2024). The engaging nature of these platforms creates a conducive environment where learners are not only entertained but also actively involved in developing and demonstrating their skills. This synergy between immersion and active participation amplifies the effectiveness of stealth assessment, as it seamlessly integrates evaluation with authentic learning experiences. Moreover, by embedding assessment items within gameplay and contextualizing them to mirror real-world challenges, stealth assessment allows students to showcase their skills, behaviors, and competencies in a manner that is both meaningful and reflective of practical applications (Shute et al. 2016a).

Building on the affordances of video games and game-based learning platforms, while simultaneously expanding possibilities for interaction, collaboration, and realism, the metaverse represents a promising frontier for implementing stealth assessment. The metaverse can be conceptualized as a networked constellation of immersive and synchronous virtual environments that afford authentic task-based interactions. As a shared digital ecosystem (Garcia et al., 2023; Iqbal & Campbell, 2023), it constitutes an assessment-rich performance space capable of unobtrusively capturing multimodal telemetry of learner actions, decisions, and strategies in real time. These affordances enable stealth assessments to connect observable interaction patterns to latent constructs of competence. Although the metaverse shares with other interactive environments the capacity to deliver immersive and engaging experiences, it distinguishes itself through its persistent and interconnected nature. Rather than being confined to discrete sessions, interactions within the metaverse unfold continuously across sustained digital spaces. Within this ecosystem, games are intrin-

sic components that draw users to facilitate exploration and interaction (Mohammed et al., 2024). Yet unlike traditional video games, which are typically organized around predefined narratives and objectives (Arif et al., 2024), the metaverse affords open-ended participation. Users can create content, cultivate social relationships, and engage in activities that mirror real-world complexities (Miller et al., 2024; Uddin et al., 2023). In turn, the metaverse supports the embedding of assessment tasks that reflect real-world complexities and provides opportunities for evaluating competencies such as problem-solving, collaboration, and adaptability in authentic contexts (Chen et al., 2023; Sidhu et al., 2024; Singh et al., 2024). Despite its considerable promise, the integration of stealth assessment into metaverse environments remains largely unexplored and represents a gap in current research.

1.3 Research gap and questions

Despite growing recognition of stealth assessment as a valid approach to evaluating learning, its application remains largely confined to game-based and simulation contexts, with relatively little exploration in metaverse environments. Existing studies have tended to emphasize competencies such as creativity, problem-solving, and collaboration (e.g., Shute and Rahimi 2021; Shute et al. 2016b), but far less attention has been given to technical competencies that underpin employability in technology-driven economies. This imbalance is problematic, given that technical skills are increasingly identified as core requirements for workforce readiness (World Economic Forum, 2025). The absence of empirical research linking stealth assessment, metaverse environments, and technical skills training also leaves a significant gap in both theory and practice. Guided by these gaps, the present study is structured around the following questions:

- RQ1. How effective is stealth assessment in measuring technical competencies in a metaverse-based training environment?
- RQ2. What is the difference in technical skill development between metaverse-based stealth assessment and traditional assessment methods?
- RQ3. To what extent can technical skills acquired in metaverse training be transferred to real-world tasks?

To address the research gaps and answer the research questions, this study is structured as a two-part investigation. Study 1 examines the effectiveness of stealth assessment in measuring technical skills within a metaverse environment (RQ1). This initial phase seeks to establish the validity and reliability of stealth assessment for technical education. Establishing reliability and effectiveness in this context is a critical first step before expanding to broader evaluations. Study 2 builds on these findings by exploring the practical applications of metaverse-based PC-building simulations. Specifically, it investigates their impact on learning outcomes compared to traditional assessment approaches (RQ2) and evaluates the extent to which competencies developed in the metaverse transfer to real-world tasks (RQ3). This two-part investigation is significant because it demonstrates how stealth assessment within metaverse-based technical training can extend theoretical foundations, provide practical guidance for

educators, and inform workforce and policy initiatives by linking digital skill development to real-world performance.

2 Evidence-centered design

2.1 Competency model

Adhering to the principles of ECD (Mislevy et al., 2003), an important step in the development process was to construct and validate competency, evidence, and task models. To develop the competency model, three instructors teaching PC building in undergraduate courses were invited to participate in a structured brainstorming session. Consulting domain experts to identify competencies of interest is a common approach in competency model development (Rahimi et al., 2024). During the session, the instructors shared their expertise and insights into the key skills and knowledge required for effective PC building. They discussed common student challenges, compared teaching strategies, and evaluated which skills should be emphasized to reflect authentic practices in the field. The brainstorming process was guided by the metaverse's learning objectives and mechanics, which emphasized building functional PCs in a simulated environment under constraints like budget, specifications, and troubleshooting scenarios. Through this collaborative process, the group identified four key competency domains for the PC Building Competency Model: (1) foundational PC knowledge, (2) strategic decision-making, (3) system assembly proficiency, and (4) diagnostic and troubleshooting. Each competency was carefully defined, justified based on the requirements of effective PC building, and linked to specific game mechanics within the metaverse simulator. These competencies serve as the foundation for evidence, task, and assembly models, ensuring that the simulator's design aligns with both learning objectives and gameplay experiences. Table 1 provides an overview of the competencies, their definitions, and the rationale for their inclusion.

2.2 Evidence model

After coming up with the final competency model, the group collaborated to build the evidence models, which define how observable actions in the metaverse-based training demonstrate the competencies identified. The group identified key observable indicators for each competency, such as correct component selection for foundational PC knowledge, budget adherence for strategic decision-making, and accurate troubleshooting paths for diagnostic expertise. For example, the evidence model for foundational PC knowledge focuses on observable actions that reflect a player's understanding of component roles, compatibility, and specifications. Specifically, evidence is gathered by tracking whether players correctly identify the purpose of components (e.g., selecting a GPU for graphics rendering) and choose compatible parts (e.g., ensuring the CPU matches the motherboard socket type). The evidence model for system assembly proficiency assesses a player's ability to accurately and

Table 1 PC Building Competencies with Definitions and Rationale

Competency	Definition	Why it was included
Foundational PC Knowledge	The ability to understand the purpose, functions, and specifications of PC components and their compatibility requirements.	Building a PC requires a foundational understanding of the function of each component and their compatibility (e.g., motherboard socket type). Without this knowledge, learners cannot make informed decisions about component selection.
System Assembly Proficiency	The ability to accurately assemble PC components by following correct procedures and ensuring proper connections.	Correctly assembling the hardware components is critical to ensuring the PC functions properly. This process involves following the proper sequence and ensuring accurate placement of PC parts.
Diagnostic and Troubleshooting	The ability to identify and resolve issues in PC builds by systematically diagnosing problems and applying effective solutions.	Diagnosing and systematically fixing errors is an essential technical skill for PC builders, whether it is resolving hardware compatibility issues or addressing problems like a system failing to boot.
Strategic Decision-Making	The ability to evaluate constraints (e.g., budget, user preferences) and make optimal choices to achieve desired outcomes in PC builds.	Building a PC often involves making trade-offs between cost and performance or adapting to constraints like budget or user needs. This competency reflects the player's ability to prioritize and balance multiple factors.

efficiently assemble a virtual PC by following the proper sequence of steps, ensuring correct placement of components, and making the necessary connections to create a functional system. Observable actions include adherence to the correct order of installation (e.g., installing the CPU before attaching the cooler), accurate placement of components (e.g., ensuring the GPU is securely inserted into the PCIe slot), and proper setup of connections (e.g., connecting the power supply to the motherboard and storage devices). Errors such as skipping critical steps or misplacing components are flagged during the simulation, and scoring is based on the completion rate,

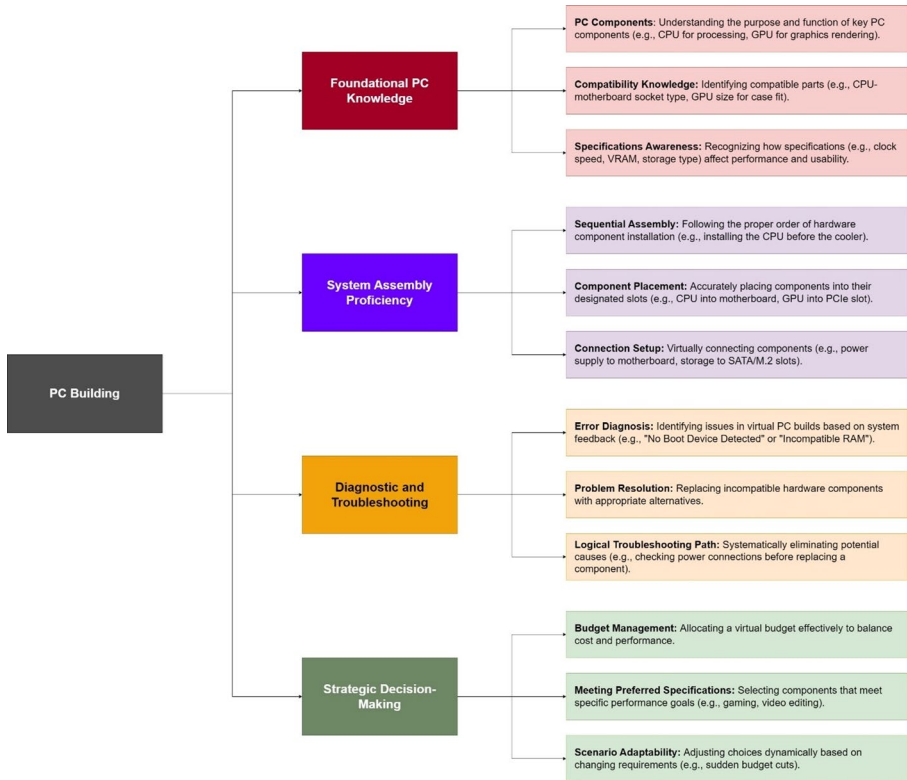


Fig. 2 PC Building Competency Model with Example Observables

order adherence, and the number of corrective attempts required. Figure 2 shows the PC Building Competency Model with examples of observable actions for each competency.

2.3 Task model

With the evidence models defined, the team proceeded to design task models to elicit observable evidence for each competency. These task models were structured to create realistic PC-building scenarios that align with the game's learning objectives and mechanics. The tasks were designed to challenge players in various areas, including component selection, strategic decision-making, system assembly, and troubleshooting. Each task was explicitly tied to the competencies and crafted to ensure a progression of difficulty, allowing players to develop their skills incrementally. Each task incorporated observable actions tied to specific scoring criteria, such as time taken, number of errors, and accuracy of decisions or solutions. These tasks provided players with opportunities to demonstrate their competencies while receiving real-time feedback to support learning and improvement. A comprehensive overview of the tasks, including example scenarios, scoring criteria, and difficulty levels, is provided in Appendix A.

2.4 Computational scoring pipeline

The stealth assessment system operationalized the evidence models through automated event logging and rule-based scoring algorithms. Each player action (e.g., component selection, assembly order, troubleshooting step) was recorded in log files and matched against evidence rules that specified success, error, or correction. Actions were weighted according to their diagnostic value. For example, correctly matching CPU–motherboard compatibility contributed positively to Foundational Knowledge, while misplacements in System Assembly triggered error codes until corrected. In addition to correctness, the system also considered efficiency and recovery. Time-stamped logs enabled scoring of task completion speed, number of corrective attempts, and sequence adherence. For example, skipping a step in assembly reduced the sequence accuracy score, while correcting the mistake within minimal retries partially recovered points. Similarly, troubleshooting scores reflected both whether the error was resolved and the logical sequence of steps taken (i.e., penalizing redundant or random attempts). Weighted indicators were then aggregated into continuous scores for each competency domain. Each competency score ranged from 0 to 100, with thresholds aligned to the categorical performance levels defined in Appendix A. This computational pipeline ensured that assessment outcomes were reproducible, transparent, and directly tied to observable in-game behaviors.

2.5 Metaverse application

A key component of this study is the metaverse application in which the stealth assessments are embedded. Developing an entire virtual world is a complex and resource-intensive task. Thus, utilizing an existing platform is a logical approach. Fortunately, the institution hosting this study features an established educational metaverse called the MILES Virtual World (Garcia et al., 2023). This platform provides a versatile foundation for integrating learning tools and assessments into an immersive environment. It has also been previously employed in educational research, underscoring its reliability and pedagogical suitability as a platform for studying learning processes in virtual environments (Garcia, 2025). Initially designed as a digital school environment for social interactions (Garcia et al., 2024), MILES Virtual World has undergone significant updates to support immersive learning experiences. Its latest iteration, the PC Building Edition (Fig. 3), marks the first attempt to transform MILES Virtual World into a learning platform. In this version, students can simulate building personal computers by assembling various hardware components. It introduces challenges such as constructing a PC within a specific budget (Fig. 4) or designing a PC build tailored to certain specifications (e.g., a high-performance gaming rig). Adding a PC building simulator in the metaverse offers a risk-free environment where students can experiment with configurations without the fear of damaging expensive equipment. The primary objective of the PC Building Edition is to eliminate the reliance on physical hardware while maintaining a hands-on, immersive learning experience.



Fig. 3 MILES Virtual World: PC Building Edition



Fig. 4 Budget-Constrained PC Building Challenge with Required Components Checklist

3 Study 1: stealth assessment in the metaverse

3.1 Methods

3.1.1 Data analysis

The data collected were analyzed to evaluate the effectiveness of the stealth assessment system in measuring PC-building competencies. Descriptive statistics were used to summarize participants' performance that are categorized by scoring lev-

els and difficulty levels. Between-group differences (novices vs. experienced participants) for each competency domain were examined using independent-samples *t*-tests. Effect sizes were computed using Cohen's *d* to quantify the magnitude of differences, with benchmarks of 0.2, 0.5, and 0.8 representing small, medium, and large effects, respectively. In addition, explained variance (r^2) was reported for correlation analyses, and semi-partial r^2 (sr^2) was reported for regression models to indicate unique contributions of each predictor. Adjusted R^2 was also calculated for overall model fit. Reliability analyses included two approaches. First, the internal consistency of the stealth assessment metrics was evaluated using Cronbach's alpha (α). Second, intraclass correlation coefficients (ICC) were computed in two contexts: (a) a two-way random ICC to assess inter-rater reliability among expert evaluators and (b) a two-way mixed ICC (absolute agreement) to examine the degree of alignment between stealth assessment metrics and expert evaluations. Validity analyses were then conducted. Pearson correlation analysis was conducted to examine the relationships between stealth assessment metrics and aggregated expert evaluation scores. Multiple linear regression was employed to predict expert scores using stealth assessment metrics as independent variables, assessing the predictive validity of the system. Statistical analyses were performed separately for novice and experienced groups to investigate differences in the accuracy and reliability of the stealth assessment system across varying expertise levels.

3.1.2 Participants and procedures

The study recruited 40 participants from computing and engineering programs, categorized into novices (limited or no prior experience with PC assembly or troubleshooting) and experienced individuals (at least 6 months of experience in building or maintaining PCs) based on self-reported expertise. Expertise classification was based on self-reported experience duration and prior exposure to PC assembly or troubleshooting tasks. Participation was voluntary, and all individuals provided informed consent. On September 16, 2024, participants visited a controlled laboratory equipped with desktop computers running the metaverse application. After a 15-minute briefing on study objectives, procedures, and the metaverse platform, participants completed a practice session to familiarize themselves with game mechanics, controls, and in-game resources. They then had one hour to complete two random PC-building scenarios: (1) a “*Basic Build*,” which involved assembling a functional PC while focusing on component roles, compatibility, and sequential assembly, and (2) a “*Troubleshooting Task*,” which required diagnosing and resolving three pre-configured issues (Appendix B). The metaverse's stealth assessment system tracked targeted metrics in real-time, while expert evaluators later reviewed participants' virtual builds and troubleshooting solutions using a standardized rubric (Appendix C) aligned with the PC-building competency model. All procedures adhered to ethical guidelines, ensuring participant data was anonymized and used solely for research purposes.

3.2 Results

3.2.1 Reliability of stealth assessment and expert evaluations

Cronbach's alpha for the stealth assessment metrics was 0.84, indicating strong internal consistency across tasks. Separate analyses by expertise level also showed high reliability, with novices at $\alpha=0.82$ and experienced participants at $\alpha=0.87$. These results confirm that the system produced stable measurements across groups. The rubric-based expert evaluations likewise demonstrated strong inter-rater reliability, with an overall ICC of 0.86 (95% CI=0.78–0.91). Subdomain analyses showed consistently high agreement among raters: Foundational PC Knowledge (ICC=0.83), System Assembly Proficiency (ICC=0.88), Diagnostic and Troubleshooting (ICC=0.85), and Strategic Decision-Making (ICC=0.84). These findings confirm that the expert rubric yielded reliable evaluations across domains.

Participants' performance on the PC-building tasks varied across competencies, difficulty levels, and expertise levels (Table 2). Independent-samples t-tests confirmed that experienced participants consistently outperformed novices across all competencies, with all differences reaching statistical significance and accompanied by large effect sizes. For Foundational PC Knowledge, novices scored 83.5 ± 6.3 compared to 91.2 ± 4.1 for experienced participants ($t = -4.58$, $p < .001$, $d = 1.45$). In System Assembly Proficiency, novices averaged 87.4 ± 7.2 while experienced participants reached 94.6 ± 4.5 ($t = -3.79$, $p < .001$, $d = 1.20$). The largest gap appeared in Diagnostic and Troubleshooting, where novices (70.8 ± 10.4) scored lower than experienced participants (82.3 ± 7.8) on error identification and resolution tasks ($t = -3.96$, $p < .001$, $d = 1.25$). Finally, in Strategic Decision-Making, novices (72.8 ± 11.5) performed significantly lower than experienced peers (81.5 ± 9.0) when making trade-offs under constraints ($t = -2.66$, $p = .011$, $d = 0.84$). The observed significant differences between novices and experienced participants were expected, given the differing levels of prior exposure to PC building. These results provide evidence of discriminant validity, as the stealth assessment system successfully distinguished between groups with theoretically different competency levels.

3.2.2 System–expert agreement by expertise level

Pearson correlation and ICC analyses were conducted to examine differences in the alignment between stealth assessment metrics and expert evaluations (Table 3). Both groups exhibited strong correlations across competencies, with experienced participants showing slightly higher correlations overall. For novices, the strongest cor-

Table 2 Performance Descriptives and Group Comparisons Across Competencies

Competency	Overall Mean $\pm$ SD	Novices Mean $\pm$ SD	Experienced Mean $\pm$ SD	t	p	d
Foundational PC Knowledge	87.4 ± 5.8	83.5 ± 6.3	91.2 ± 4.1	-4.58	<0.001	1.45
System Assembly Proficiency	91.2 ± 6.1	87.4 ± 7.2	94.6 ± 4.5	-3.79	<0.001	1.20
Diagnostic and Troubleshooting	76.5 ± 9.1	70.8 ± 10.4	82.3 ± 7.8	-3.96	<0.001	1.25
Strategic Decision-Making	79.6 ± 10.5	72.8 ± 11.5	81.5 ± 9.0	-2.66	0.011	0.84

Table 3 Agreement Between Stealth Assessment and Expert Evaluations by Expertise Level

Competency	Group	<i>r</i>	<i>r</i> ²	<i>p</i>	ICC (95% CI)
Foundational PC Knowledge	Novices	0.68	46%	0.023	0.75 (0.67–0.83)
	Experienced	0.72	52%	0.015	0.81 (0.73–0.89)
System Assembly Proficiency	Novices	0.74	55%	0.012	0.78 (0.70–0.86)
	Experienced	0.78	61%	0.008	0.85 (0.77–0.91)
Diagnostic and Troubleshooting	Novices	0.69	48%	0.030	0.73 (0.65–0.81)
	Experienced	0.80	64%	0.010	0.84 (0.76–0.90)
Strategic Decision-Making	Novices	0.62	38%	0.045	0.71 (0.63–0.79)
	Experienced	0.68	46%	0.022	0.78 (0.70–0.86)

relation was observed in System Assembly Proficiency ($r=.74$; $p=.012$), while for experienced participants, Diagnostic and Troubleshooting showed the highest alignment ($r=.80$; $p=.010$). ICC values were also higher for experienced participants ($ICC=0.84$) compared to novices ($ICC=0.77$).

3.2.3 Predictive validity of stealth assessment metrics

The results of regression analyses revealed differences in the predictive power of stealth assessment metrics (Table 4). For novices, the model explained 49% of the variance in expert scores (adjusted $R^2 = 0.49$), with System Assembly Proficiency ($\beta=0.48$; $p=.014$; $sr^2=0.23$) and Foundational PC Knowledge ($\beta=0.35$; $p=.002$; $sr^2=0.18$) emerging as the strongest predictors. For experienced participants, the model accounted for 54% of the variance (adjusted $R^2 = 0.54$), with Diagnostic and Troubleshooting ($\beta=0.51$; $p=.012$; $sr^2=0.26$) and Strategic Decision-Making ($\beta=0.42$; $p=.003$; $sr^2=0.21$) contributing significantly to expert scores. Despite these differences by expertise level, the regression models for both groups demonstrated strong predictive validity. Furthermore, it is worth noting that some beta values for novices were higher than those for experienced participants, such as System Assembly Proficiency ($\beta_{\text{novices}}=0.48$, $\beta_{\text{Experienced}}=0.36$). This unexpected pattern may reflect a methodological reason, such as the possibility that novices show more pronounced gains or stronger relationships in foundational competencies as they acquire new skills,

Table 4 Regression Analysis by Expertise Level

Competency	Group	β	SE	p	sr^2
Foundational PC Knowledge	Novices	0.35	0.09	0.002	0.18
	Experienced	0.32	0.10	0.005	0.15
System Assembly Proficiency	Novices	0.48	0.08	0.014	0.23
	Experienced	0.36	0.09	0.004	0.17
Diagnostic and Troubleshooting	Novices	0.30	0.11	0.018	0.12
	Experienced	0.51	0.08	0.012	0.26
Strategic Decision-Making	Novices	0.21	0.10	0.045	0.09
	Experienced	0.42	0.09	0.003	0.21

whereas experienced participants may display more nuanced performance that is less strongly correlated with specific metrics.

4 Study 2: technical skills and transferability

4.1 Methods

4.1.1 Participants and procedures

This study recruited two sections of a Computer Hardware Fundamentals class, each consisting of 40 students ($n=80$). These sections were randomly assigned to two groups using a computer-generated random number procedure to ensure that each section had an equal probability of being placed in either group. The Stealth Assessment Group (SAG) engaged in a metaverse-based PC-building game with embedded stealth assessment and adaptive feedback, while the Traditional Assessment Group (TAG) learned through conventional methods, such as instructional videos, text-based guides, and quizzes. Efforts were made to balance prior knowledge across groups, with participants self-reporting their PC-building experience. Approximately half of the participants ($n=38$) identified as novices (little or no prior experience), while the remainder ($n=42$) reported moderate or extensive experience in PC assembly and troubleshooting. To confirm baseline equivalence, independent-samples t -tests were conducted on pre-test scores across the four targeted competencies. No significant differences were observed between SAG and TAG (all $p > .05$), which indicates that both group of students started with comparable levels of PC-building knowledge and skills prior to the intervention. Both sections were taught by the same instructor, following a standardized syllabus and pacing guide to minimize instructor-related variability. To further control for potential instructor effects, the delivery of lectures, demonstrations, and feedback was scripted and applied consistently across groups. The primary difference between groups lay in the assessment method.

The experiment (Fig. 5) commenced during the first week of the trimester (January 2025) and lasted five days, with sessions of two hours per day. From Day 1 to Day 4, students engaged in structured activities and completed PC-building scenarios

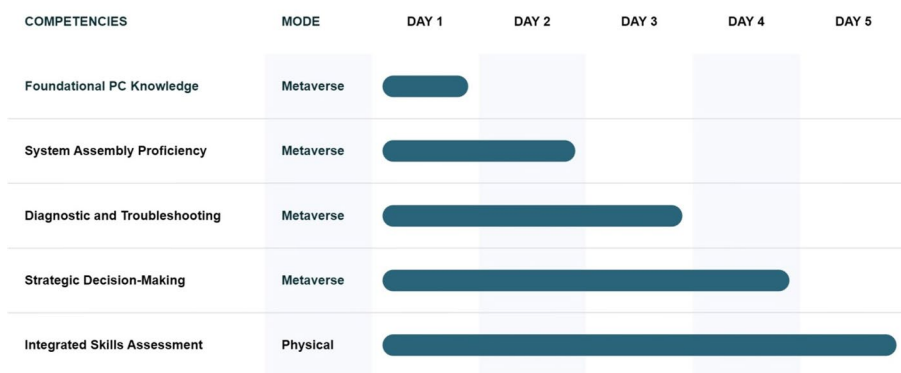


Fig. 5 Daily Distribution of Targeted Competencies in the Metaverse-Based PC-Building Training

Table 5 Pre-Test and Post-Test Scores for Each Group Across Four Days

Day	Competency	Group	Pre-Test Mean $\pm$ SD	Post-Test Mean $\pm$ SD	<i>p</i>	<i>d</i>
1	Foundational PC Knowledge	SAG	18.4 $\pm$ 2.1	19.5 $\pm$ 3.4	0.073	0.35
		TAG	15.3 $\pm$ 3.8	19.8 $\pm$ 5.1	<0.001	1.01
2	System Assembly Proficiency	SAG	14.1 $\pm$ 2.3	23.0 $\pm$ 2.9	<0.001	1.45
		TAG	19.9 $\pm$ 2.6	20.8 $\pm$ 3.0	0.159	0.34
3	Diagnostic and Troubleshooting	SAG	15.2 $\pm$ 2.4	25.4 $\pm$ 2.1	<0.001	1.85
		TAG	14.8 $\pm$ 2.5	22.1 $\pm$ 2.7	<0.001	1.20
4	Strategic Decision-Making	SAG	16.0 $\pm$ 2.2	26.0 $\pm$ 1.9	<0.001	1.95
		TAG	15.5 $\pm$ 2.3	23.0 $\pm$ 2.4	<0.001	1.40

designed to target specific competencies each day. Each session began with a pre-test to assess baseline knowledge and skills related to the day's competencies and concluded with a post-test to measure learning gains. On Day 5, students from the SAG participated in hands-on testing to evaluate the real-world transferability of skills developed in the virtual environment. They were tasked with assembling a physical PC based on a provided scenario (e.g., gaming or video editing) and troubleshooting a pre-configured issue in a non-functional system. Expert evaluators used standardized rubric aligned with the PC-building competency model to assess component placement, connection setup, and troubleshooting success. Ethical guidelines were strictly followed throughout the study, with informed consent obtained from all participants, and anonymity of data assured.

4.1.2 Data analysis

The data collected were analyzed to compare the learning outcomes of metaverse-based and traditional methods, as well as the real-world transferability of skills developed in the virtual environment. Descriptive statistics were computed for pre-test and post-test scores for both groups to summarize overall performance across the four days of training. Paired *t*-tests were performed within each group to compare daily

pre-test and post-test scores, with effect sizes (Cohen's d) computed to quantify the magnitude of learning gains. Independent t -tests were conducted to compare post-test scores between the two groups, evaluating the relative effectiveness of the two instructional methods. Metrics from virtual PC-building tasks were also analyzed using independent t -tests to assess the effectiveness of stealth assessment in promoting task-specific competencies. To evaluate the real-world transferability of skills, Pearson correlations were computed between virtual task performance metrics and hands-on PC assembly scores for the SAG, with variance explained (r^2) reported to indicate predictive strength. Additionally, where applicable, adjusted R^2 values from regression models predicting hands-on scores from virtual metrics were examined to assess overall predictive validity.

4.2 Results

4.2.1 Learning outcomes from pre-test and post-test scores

Descriptive statistics and paired t -tests were conducted to analyze learning improvements within each group (Table 5). Both groups demonstrated learning gains over the four days of training. However, the magnitude of improvement varied, and differences were not always statistically significant. While both groups showed performance improvements in foundational PC knowledge and system assembly proficiency, no statistically significant differences were observed between their pre-test and post-test scores on Day 1 for the SAG ($p = .073$) and on Day 2 for the TAG ($p = .159$). Effect sizes indicated that gains were especially pronounced for the SAG in advanced competencies, with very large effects on Day 3 ($d = 1.85$) and Day 4 ($d = 1.95$), compared to more modest effects in the TAG (Day 3, $d = 1.20$; Day 4, $d = 1.40$). These results suggest that stealth assessment was particularly effective in supporting the acquisition of higher-order competencies such as troubleshooting and decision-making.

4.2.2 Comparison of post-test scores between groups

As shown in Table 6, independent t -tests revealed significant differences between SAG and TAG on Days 3 and 4, where SAG demonstrated superior performance in diagnostic and troubleshooting compared to TAG ($p < 0.001$; $d = 1.36$) and strategic decision-making ($p < 0.001$; $d = 1.39$). However, no significant differences

Table 6 Comparison of Post-Test Scores Between Groups

Day	Competency	Group	Post-Test Mean $\pm$ SD	p	d
1	Foundational PC Knowledge	SAG	19.5 $\pm$ 3.4	0.684	0.09
		TAG	19.8 $\pm$ 3.1		
2	System Assembly Proficiency	SAG	23.0 $\pm$ 2.9	0.761	0.07
		TAG	22.8 $\pm$ 3.0		
3	Diagnostic and Troubleshooting	SAG	25.4 $\pm$ 2.1	<0.001	1.36
		TAG	22.1 $\pm$ 2.7		
4	Strategic Decision-Making	SAG	26.0 $\pm$ 1.9	<0.001	1.39
		TAG	23.0 $\pm$ 2.4		

were observed on Days 1 and 2, indicating that for foundational skills such as PC knowledge ($p = .684$; $d = 0.09$) and system assembly proficiency ($p = .761$; $d = 0.07$), both instructional methods were similarly effective. For instance, SAG (19.5 ± 3.4) and TAG (19.8 ± 3.1) achieved comparable post-test scores, suggesting equivalent impact in teaching foundational concepts. These results highlight the stronger impact of stealth assessment on advanced competencies, particularly diagnostic, troubleshooting, and decision-making skills. The observed large effect sizes on Days 3 and 4 reinforce that the metaverse-based training was especially impactful for advanced competencies.

4.2.3 Analysis of real-world transferability of skills

The results indicated strong correlations between virtual and hands-on performance for some metrics, particularly in assembly accuracy ($r = .82$; $p = .005$) and error diagnosis ($r = .79$; $p = .006$), suggesting successful transfer of these skills from the virtual to the physical environment (Table 7). Variance explained (r^2) confirmed this pat-

Table 7 Hands-On Task Performance and Correlations with Virtual Metrics

Metrics	Mean $\pm$ SD	r	r^2	p	Interpretation
Component Roles	89.3 $\pm$ 6.1	0.68	0.46	0.018	Moderate correlation, skills transfer
Compatibility Knowledge	86.7 $\pm$ 5.9	0.41	0.17	0.045	Weak correlation, partial transfer
Specifications Awareness	84.5 $\pm$ 7.0	0.34	0.12	0.071	Not significant, limited transfer
Sequential Assembly	92.1 $\pm$ 4.8	0.82	0.67	0.005	Strong correlation, excellent transfer
Component Placement	88.4 $\pm$ 5.7	0.76	0.58	0.009	Strong correlation, skills transfer
Connection Setup	85.9 $\pm$ 6.2	0.52	0.27	0.031	Moderate correlation, partial transfer
Error Diagnosis	87.8 $\pm$ 5.5	0.79	0.62	0.006	Strong correlation, excellent transfer
Problem Resolution	83.5 $\pm$ 6.8	0.47	0.22	0.042	Weak correlation, partial transfer
Logical Troubleshooting Path	81.9 $\pm$ 7.3	0.38	0.14	0.083	Not significant, limited transfer
Budget Management	78.6 $\pm$ 8.0	0.20	0.04	0.211	No significant correlation
Specification Matching	82.3 $\pm$ 7.5	0.43	0.18	0.052	Weak correlation, borderline significant
Scenario Adaptability	80.4 $\pm$ 7.8	0.31	0.10	0.097	Not significant, limited transfer

tern, with sequential assembly ($r^2 = 0.67$) and error diagnosis ($r^2 = 0.62$) indicating that more than 60% of variance in hands-on performance could be predicted from virtual task performance. For example, students demonstrated high hands-on scores for sequential assembly (92.1 ± 4.8) and component placement (88.4 ± 5.7). However, no significant correlation was observed for budget management ($r = .20$; $p = .211$) and scenario adaptability ($r = .31$; $p = .097$), which may indicate challenges in adapting abstract and context-dependent skills to physical tasks. These results suggest that while procedural and diagnostic skills transferred robustly, strategic decision-making competencies were less easily generalized from virtual to real-world contexts.

5 Discussion, implications, and limitations

Developing hands-on technical skills has traditionally relied on physical practice, but emerging metaverse environments are redefining how such competencies can be learned and assessed. In this study, the metaverse platform was paired with a stealth assessment approach to enable unobtrusive tracking of participant performance during task execution. Participants demonstrated measurable improvements in PC-building competencies, with gains evident across both assembly and troubleshooting tasks, as well as in real-world transfer outcomes following metaverse-based training. These findings suggest that simulation-based environments, when coupled with embedded assessment mechanisms, can facilitate applied learning while enabling its observation in computing education contexts. While both the metaverse environment and the stealth assessment system were integral to the study, they serve distinct roles. The metaverse platform functions as the instructional medium, providing an interactive space in which participants engage in PC-building tasks and develop relevant competencies. In contrast, the stealth assessment system operates as a measurement mechanism, unobtrusively capturing behavioral data such as task sequences, errors, and completion patterns. As such, the observed performance improvements are primarily attributable to the instructional affordances of the metaverse environment, while the contribution of stealth assessment is the enhancement of the granularity and validity of performance measurement.

5.1 Stealth assessment applications in a metaverse environment

To date, the metaverse remains an untapped digital environment for conducting stealth assessments. While prior works have explored the use of the metaverse in education, these studies have primarily relied on self-reported surveys, observational studies, or traditional test-based assessments rather than leveraging in-game performance data for real-time evaluation (Onu et al., 2024; Qian et al., 2023). While valuable, such methods often fail to capture the depth and authenticity of learner interactions within the metaverse. This observation echoes the critique of Shute et al. (2017) that standardized assessments in immersive environments lack the depth needed to evaluate the richness of student learning experiences. Building on this limitation, this study represents an early attempt to apply stealth assessment in a metaverse environment. The findings suggest that stealth assessment is likewise highly suitable for meta-

verse environments, as evidenced by its ability to measure technical competencies with validity and reliability. This is consistent with earlier studies on stealth assessment in video games (e.g., Abbasi & Kazi, 2020; Fang et al., 2023; Shute & Rahimi, 2021), which emphasizes the capacity of interactive media to unobtrusively evaluate skills and competencies in engaging and authentic contexts. By extending this discreet performance-based approach to the metaverse, the study supports and strengthens the assertion of Rahimi and Shute (2024) that stealth assessment is an effective evaluation tool in technology-rich environments. Overall, the findings illustrate that stealth assessment can be adapted to measure learner behaviors in metaverse environments, which may serve as a useful medium for advancing educational assessment approaches.

5.2 Metaverse as a superior medium for technical skill development

Consistent with earlier experimental research (e.g., Vate-U-Lan & Cahill, 2024), this study provides evidence that metaverse-based approaches may offer advantages over conventional methods for certain competencies. A recurring theme in the educational metaverse literature is the metaverse's provision of realistic and engaging hands-on experiences as a core benefit over conventional platforms (Onu et al., 2024). Unlike traditional methods, which often rely on static resources like manuals or videos, the metaverse provides an interactive, 3D environment where learners can engage in authentic, task-oriented scenarios (Garcia et al., 2024; Miller et al., 2024; Ng, 2022). For example, MILES Virtual World allows students to virtually assemble a custom PC by selecting and positioning components while receiving real-time feedback on compatibility and performance issues. This hands-on approach mimics real-world assembly processes, allowing learners to practice and refine their skills without the financial or logistical risks associated with physical hardware. Such an approach is particularly advantageous in situations where face-to-face instruction is limited by resource constraints (Rudolphi-Solero et al., 2025; Tong et al., 2024), such as the availability of hardware, safety concerns, or scalability challenges. Unlike metaverse-based environments, in-person training often requires significant financial investment in equipment and infrastructure, can be limited to small groups due to space and resource constraints, and may lack the ability to provide immediate, adaptive feedback tailored to individual learners' needs. These limitations suggest that metaverse environments may provide a practical alternative platform for technical training that would otherwise be difficult or impractical in traditional face-to-face settings (Popov et al., 2024). Similar approaches have been explored in other fields, such as surgical education (Ammendola et al., 2024) and corporate training (Garg et al., 2025), where metaverse-based simulations are being tested for complex procedures.

The scenario-driven environment offered by the metaverse for hands-on training finds its roots in the principles of simulation-based education, which has long been recognized as an effective method for knowledge acquisition and skill development across various disciplines (Vlachopoulos & Makri, 2017). During simulation-based training in a controlled virtual space, students benefit from the safety of an environment where they can experiment, make mistakes, and refine their skills without facing the consequences of real-world errors. In disciplines like medical education,

metaverse-based simulation has been shown to accelerate the mastery of technical skills compared to traditional training methods (Popov et al., 2024). One explanation is the capability of the metaverse to seamlessly bridge the gap between theoretical knowledge and practical application. For example, in the MILES Virtual World, students troubleshooting a virtual PC are presented with a scenario where the system fails to boot. They must identify potential issues (e.g., incompatible RAM, missing storage devices, or improperly installed GPUs) by analyzing system error messages, reviewing component specifications, and simulating the replacement or adjustment of parts. This process mirrors real-world troubleshooting tasks while offering the added benefits of experimenting with multiple configurations and resolving complex issues in a risk-free environment that fosters exploration and encourages learning from mistakes. Drawing from research on interactive simulations and immersive learning environments (Hammouda et al., 2025; Miller et al., 2024), manipulating virtual PC components and observing their outcomes in an engaging, lifelike environment fosters deeper understanding and skill retention. Overall, these results demonstrate the metaverse's advantages in replicating real-world contexts while offering flexibility and interactivity to enhance learning experiences.

5.3 Boundary conditions of metaverse effectiveness

An interesting and somewhat unexpected result was that traditional instructional methods performed equally well as the metaverse-based approach in developing foundational competencies such as basic PC knowledge and system assembly proficiency. This outcome may be explained by the nature of these competencies, which rely heavily on declarative knowledge (e.g., recognizing component roles and specifications) and procedural repetition (e.g., following a fixed sequence of assembly). Prior research has shown that such lower-order skills can often be acquired effectively through conventional resources such as manuals, videos, and step-by-step guides (Vlachopoulos & Makri, 2017). In contrast, the distinctive affordances of the metaverse (e.g., interactivity, adaptive feedback, and scenario-based challenges) appear to yield greater benefits in domains that demand diagnostic reasoning, decision-making under constraints, and adaptive problem-solving (Hunt et al., 2023; Petersen et al., 2022). This finding contributes to ongoing discussions on blended instructional design, suggesting that while immersive environments may be superior for cultivating higher-order and context-dependent skills, traditional approaches remain pedagogically sufficient and resource-efficient for teaching foundational knowledge (Conrad et al., 2024; Ryan et al., 2022; Santilli et al., 2025; Yang et al., 2024). Future research should examine how hybrid models can strategically align instructional modality with skill complexity.

5.4 Real-world transferability of skills developed in the metaverse

Interestingly, students demonstrated a clear ability to apply skills developed in the metaverse-based PC-building environment to real-world tasks. These findings provide empirical evidence supporting the successful transfer of skills acquired in the metaverse to analogous scenarios in physical contexts. This is a notable strength of

the study, as most experimental research on the metaverse tends to focus on comparing methods without evaluating the practical application of the skills acquired (e.g., Vate-U-Lan & Cahill, 2024). Research in similar technology-enhanced learning environments provides insights into this phenomenon. For instance, Dobrowolski et al. (2021) highlighted that the immersive nature of virtual reality (VR) platforms creates an engaging experience, which fosters better skill transferability and practical application. This notion is consistent with the findings of Petersen et al. (2022), who posited that immersive technologies facilitate learning through presence and agency. In augmented reality (AR) environments, Hunt et al. (2023) emphasized that participants' higher engagement levels and the perceived authenticity of the training environment were critical in enhancing their performance. These findings are equally relevant to the metaverse, as illustrated in the MILES Virtual World. For instance, students were challenged to assemble a high-end gaming PC while adhering to a strict budget. The realistic nature of the task enhanced the perceived authenticity of the training by replicating the constraints and decision-making processes learners would face in practical applications. The dynamic and interactive nature of the metaverse further fostered a strong sense of presence and agency, as students could directly manipulate virtual components and immediately observe the consequences of their choices. Building on earlier research, the heightened engagement and realism of the metaverse seem to have facilitated the development of practical skills that were confidently applied in physical PC-building tasks.

Embedded within the training, lifelike challenges, and problem-solving opportunities is the system feedback mechanism of the MILES Virtual World. This mechanism is important because immediate feedback is essential for learning and skill development in the metaverse (Damaševičius & Sidekerskienė, 2024). For example, when students encounter errors in their virtual builds (e.g., attempting to install incompatible RAM or forgetting to connect a power supply), the metaverse provided contextual feedback with detailed explanations to guide their corrections. With this feature, students are empowered to troubleshoot and resolve the issues independently. This aspect connects directly to the tenets of constructivism and self-directed learning (Lee & Ahn, 2025; Li & Yu, 2023). From a constructivist perspective, the metaverse likewise leverages the concept of an “invented reality,” where learning is deeply influenced by the learner's ability to perceive and act within the virtual environment (Aiello et al., 2012). Knowledge is not passively absorbed but actively constructed through personal experiences, mental models, and perceptual mechanisms. Whether real or simulated, these processes are intrinsically linked to the learner's interactions with their environment. Moreover, this interplay between knowledge construction and active engagement also reinforced procedural memory (Dubinsky & Hamid, 2024). For example, MILES Virtual World required students to follow precise sequences (e.g., installing standoffs before securing the motherboard) to ensure that procedural steps were practiced and internalized. The combination of immediate feedback, active engagement, and structured practice underscores the metaverse's potential as a promising platform for developing technical skills in an immersive and meaningful way.

5.5 Practical, theoretical, and policy implications

From a practical standpoint, this study demonstrates that metaverse-based environments can function as cost-effective, scalable, and pedagogically sound alternatives to traditional hardware-dependent training. In technical domains such as PC building, financial and logistical barriers often limit student access to authentic practice opportunities, as physical hardware is expensive, prone to wear, and requires specialized facilities. By contrast, immersive virtual platforms such as MILES Virtual World enable students to repeatedly practice complex assembly and troubleshooting tasks in a risk-free environment. Errors become low-stakes opportunities for learning rather than costly mistakes, which is a feature particularly valuable for TVET programs operating under resource constraints (Popov et al., 2024). In addition, the embedded feedback mechanisms within the metaverse environment ensure that learning is iterative and adaptive (Baudier et al., 2025), allowing students to refine their skills through immediate corrective guidance. Industry applications are also evident. Technology-driven organizations could integrate such environments into professional training (e.g., Adarkwah & Islam, 2025; Garg et al., 2025; Nofal et al., 2025), enabling employees to rehearse tasks such as hardware maintenance or system troubleshooting in a simulated yet authentic environment. The practical implication is therefore twofold: (1) metaverse-based training may broaden access to skill development by lowering material and financial barriers, and (2) it provides repeatable, scalable opportunities for learners and employees to gain competence through interactive practice.

Theoretically, this study advances the discourse on stealth assessment by extending its applicability to metaverse-based environments. While stealth assessment has been primarily explored in serious games and traditional digital simulations (Abbasi & Kazi, 2020; Shute & Rahimi, 2021), this research demonstrates how it can be embedded into immersive, scenario-driven environments that mirror authentic technical tasks. The application of ECD within the metaverse illustrates the effectiveness of aligning competency models, evidence models, and task models to ensure that assessment is seamlessly integrated into authentic learning activities. This integration supports a shift in assessment theory from decontextualized testing toward process-based evaluation of learner performance (Rahimi & Shute, 2024; Shute et al., 2017). Furthermore, the findings show how metaverse affordances (e.g., presence, agency; Garcia, 2026) expand the range of competencies that can be measured, from procedural accuracy to diagnostic reasoning and strategic decision-making (Chu et al., 2025; Hunt et al., 2023). Operationalizing stealth assessment in this way contributes to theory in two important ways: it illustrates that immersive platforms can capture unobtrusive evidence of learning, and it indicates that metaverse environments may be able to capture not only cognitive skills but also procedural, diagnostic, and adaptive competencies that are less frequently assessed in traditional approaches.

From a policy perspective, the findings point to emerging opportunities for integrating metaverse-based learning and assessment into TVET and higher education systems. Policymakers could consider supporting pilot programs that deploy metaverse-based training modules in resource-limited institutions, particularly where access to physical hardware is constrained. Such initiatives would align with broader policy goals to democratize access to technical education and promote digital equity

(UNESCO, 2022). However, the evidence base remains narrow as this study focused on one domain and a limited participant sample. Thus, while promising, policy recommendations must be framed as exploratory rather than prescriptive. Larger-scale, multi-domain studies are needed before widescale adoption. Ethical concerns also warrant policy-level attention. Stealth assessment relies on the continuous collection of learner performance data, raising questions of privacy, data governance, and algorithmic transparency (Damaševičius & Sidekerskienė, 2024). Policy frameworks should consequently establish clear guidelines on how learner data is collected, stored, and used to ensure that educational benefits are balanced against risks to student autonomy and privacy. Overall, these findings suggest that while metaverse-based stealth assessment may offer meaningful opportunities for enhancing technical education, its policy adoption should proceed gradually and supported by ethical safeguards.

5.6 Study limitations and future research needs

While this study provides valuable insights into the use of stealth assessment in a metaverse environment, several limitations should be acknowledged. First, the sample size was relatively small and limited to students from computing and engineering programs. Although novice–expert comparisons provided meaningful differentiation, the modest number of participants constrains the statistical power of some analyses and limits the generalizability of findings to broader learner populations. Future studies should employ larger and more diverse samples, including participants from other TVET domains and higher education contexts. Additionally, real-world transfer was assessed only for the metaverse group. As a result, direct comparisons of transfer effectiveness across instructional conditions were not possible, which limits conclusions regarding the relative impact of the metaverse on real-world performance.

Second, the study focused exclusively on a specific domain that limits the applicability of the findings to other technical skills or fields. Within the broader frameworks of TVET, PC building represents only one of several core competencies. Expanding the scope to include additional modules, diverse disciplines, or multi-skill training scenarios could provide a more comprehensive understanding of the metaverse's capabilities as a learning medium.

Third, while the metaverse environment was designed to closely mimic real-world tasks and physical campuses of the university, it may not fully capture the tactile and sensory experiences of physical PC building, such as handling hardware components or managing spatial constraints. These aspects are integral to authentic skill development in TVET programs (Peter et al., 2025), and their absence may influence the transferability of certain competencies.

Fourth, the metaverse platform used in this study was PC-based. This design decision was intentional to ensure accessibility, scalability, and consistency across participants, since not all institutions or learners have access to VR headsets or haptic devices. While this version required students to navigate the environment using standard computer peripherals, which may not fully replicate the immersive potential of VR, AR, or MR/XR-based metaverse experiences, it provided a practical and cost-effective environment for piloting stealth assessment. Future research could build

upon this foundation by integrating advanced hardware such as VR headsets or haptic devices to further enhance realism and interactivity (Liu et al., 2025).

Finally, the short duration of the training and assessment phases may not fully reflect the long-term retention of skills developed in the metaverse. Longitudinal studies situated within TVET frameworks are needed to evaluate the durability of these competencies and their continued application in real-world contexts over time. Addressing these limitations in future research will strengthen the evidence base for integrating metaverse-based learning and stealth assessment into technical education and vocational training pathways.

6 Conclusion

This two-part investigation provides empirical evidence on the suitability of stealth assessment within a metaverse context and the effectiveness of metaverse-based training in improving learning outcomes and facilitating skill transfer to real-world tasks. Validating the efficacy of stealth assessment in metaverse environments is fundamental to advancing educational technologies that align with modern pedagogical frameworks and industry needs. Beyond this validation, the research offers valuable insights into the broader application of adaptive learning technologies and contextualized assessments. It underscores the importance of aligning ECD principles with immersive environments to create authentic and meaningful learning experiences. Furthermore, the research lays the groundwork for future studies to explore metaverse-based training across other technical and multi-domain competencies, examine the long-term retention and application of acquired skills, and address critical factors such as accessibility, interoperability, inclusivity, and the responsible use of learner data. These directions resonate with ongoing efforts under UNESCO's TVET digital transformation framework and the OECD Learning Compass 2030, both of which advocate for competence-based and technology-mediated education. With significant implications for educators, researchers, and policymakers, this research contributes to the growing evidence base supporting the integration of metaverse environments and stealth assessment in technical education and workforce development.

Appendix 1

The simulator implemented a rule-based event logging system that automatically mapped player actions to competency indicators. Each in-game action (e.g., selecting a CPU–motherboard pair, inserting a GPU into a PCIe slot, replacing incompatible RAM) was logged with timestamps and evaluated against predefined evidence rules. Actions were coded as correct, incorrect, or corrective, and weighted by their diagnostic importance. For example, selecting the correct component role added to Foundational Knowledge, following the correct order of assembly steps contributed to System Assembly Proficiency, and completing logical troubleshooting steps contributed to Diagnostic and Troubleshooting. These weighted values were summed into domain-level competency scores, with performance categories (Excellent, Good,

Needs Improvement) reflecting thresholds defined by the scoring criteria in the tables below. This framework ensured that the scoring process remained both automated and directly aligned with the evidence-centered design principles.

Table 8 Sample tasks, scenarios, observables, and scoring criteria for pc building competencies.

Competency	Subcompetency	Task	Example Scenario	Observables	Scoring Criteria	Difficulty Level
Foundational PC Knowledge	Component Roles	Identify Components	Select the appropriate GPU for a high-performance gaming PC build.	Correctly identifying the component's role in the system (e.g., GPU for graphics rendering).	100% correct = Excellent; 75–99% = Good; <75% = Needs Improvement.	Beginner
	Compatibility Knowledge	Match Components	Select a CPU and motherboard with compatible socket types.	Compatibility errors flagged by the simulator.	0 errors = Excellent; 1–2 errors = Good; >2 errors = Needs Improvement.	Intermediate
	Specifications Awareness	Match Performance Needs	Select components (CPU, GPU, RAM) for a 4 K video editing build.	Components meet the required performance specifications (e.g., clock speed, VRAM).	100% of specs met = Excellent; 80–99% = Good; <80% = Needs Improvement.	Advanced
System Assembly Proficiency	Sequential Assembly	Follow Assembly Steps	Assemble a PC starting with CPU installation and ending with cable connections.	Correct order of steps followed, skipped or misordered steps flagged.	100% in sequence = Excellent; 90–99% = Good; <90% = Needs Improvement.	Beginner to Advanced
	Component Placement	Install Components	Place the CPU, GPU, RAM, and storage correctly in their respective slots.	Accurate placement of components; number of placement errors flagged.	100% correct = Excellent; 90–99% = Good; <90% = Needs Improvement.	Beginner to Intermediate
	Connection Setup	Connect Components	Establish virtual connections between the power supply, motherboard, and storage devices.	Number of correct connections made; errors flagged.	100% correct = Excellent; 90–99% = Good; <90% = Needs Improvement.	Intermediate

Table 8 (continued)

Competency	Subcompetency	Task	Example Scenario	Observables	Scoring Criteria	Difficulty Level
Diagnostic and Troubleshooting	Error Diagnosis	Identify Issues	Diagnose the cause of a PC not booting due to incompatible RAM.	Correct identification of the issue within minimal attempts.	1st attempt correct=Excellent; 2–3 attempts=Good; >3 attempts=Needs Improvement.	Intermediate to Advanced
	Problem Resolution	Fix Errors	Replace the incompatible RAM with a compatible alternative.	Correct replacement of the problematic component.	1st replacement correct=Excellent; 2–3 replacements=Good; >3=Needs Improvement.	Intermediate
	Logical Troubleshooting Path	Systematically Troubleshoot	Follow a logical sequence to resolve an issue caused by multiple errors (e.g., loose connections, incompatible RAM).	Steps followed systematically; unnecessary actions avoided.	Logical path with 0 redundant actions=Excellent; minor deviations=Good; major deviations=Needs Improvement.	Advanced
Strategic Decision-Making	Budget Management	Optimize Budget	Build a gaming PC within a PHP100,000 budget, prioritizing GPU and CPU performance.	Final cost relative to budget; performance goals achieved.	<10% under budget=Excellent; within budget=Good; over budget=Needs Improvement.	Intermediate to Advanced
	Specification Matching	Match Specs	Build a system capable of handling 3D rendering tasks with required performance benchmarks.	Components selected meet performance benchmarks.	100% specs met=Excellent; 80–99%=Good; <80%=Needs Improvement.	Intermediate to Advanced
	Scenario Adaptability	Adjust to Changes	Adapt a build after a budget cut or new requirement is introduced mid-task.	Time taken to adjust; accuracy of new build.	<2 min to adapt=Excellent; 2–4 min=Good; >4 min=Needs Improvement.	Advanced

Appendix 2

This appendix presents representative examples of the PC-building and troubleshooting tasks implemented within the metaverse environment. These scenarios illustrate the types of activities participants encountered during the study and provide additional detail on task design, objectives, and expected actions. The examples are not exhaustive but are intended to demonstrate how core competencies (e.g., component identification, system assembly, compatibility evaluation, and diagnostic reasoning) were operationalized within the virtual platform. All scenarios were designed to be feasible within a metaverse-based simulation and to generate observable performance data aligned with the study's stealth assessment framework.

Table 9 Sample PC-building and troubleshooting scenarios

B1. Example Basic Build Scenarios

Scenario	Objective	Key Requirements	Skills Assessed	Success Criteria
Entry-Level Office PC	Build a functional system for basic tasks	Compatible CPU–motherboard pairing, RAM installation, storage setup	Component identification, compatibility	System boots successfully with no errors
Mid-Range Gaming PC	Assemble a performance-oriented PC	Dedicated GPU selection, sufficient PSU, proper component matching	Assembly sequencing, compatibility decisions	Stable system with all components recognized
Budget-Constrained Build	Build within a fixed budget	Select cost-effective yet compatible components	Strategic decision-making, trade-offs	Functional system within constraints

B2. Example Troubleshooting Scenarios

Case	Description	Underlying Issue	Expected Actions	Competencies Assessed
No Boot Device Detected	System powers on but no OS loads	Missing or unassigned storage device	Install/select storage, configure boot order	Diagnostic reasoning
Incompatible RAM	System fails to start or shows error	RAM not supported by motherboard	Replace with compatible RAM	Component knowledge
Missing Power Connection	System does not power on	Unconnected motherboard or CPU power cable	Inspect and connect power cables	Systematic troubleshooting
GPU Not Detected	Display output not working	GPU not installed properly or not powered	Reseat GPU, connect PCIe power	Hardware configuration understanding
Incorrect Boot Order	System bypasses OS or loops	Boot priority misconfigured in BIOS	Adjust boot sequence	Logical problem-solving

Appendix 3

This appendix presents the standardized rubric used by expert evaluators to assess participants' performance across key PC-building competency domains. The rubric was developed based on the study's competency framework and was designed to ensure consistent, objective, and comparable evaluations of both system assembly and troubleshooting performance. Each competency domain is rated on a four-level scale ranging from Poor (1) to Exemplary (4), with descriptors outlining observable behaviors and performance quality at each level. The rubric was applied to participants' virtual outputs and actions within the metaverse environment, enabling structured evaluation aligned with the study's learning and assessment objectives.

Table 10 Standardized Rubric for PC-Building Competency Evaluation

Competency Domain	1 – Poor	2 – Developing	3 – Proficient	4 – Exemplary
Foundational PC Knowledge	Misidentifies most components; frequent compatibility errors.	Some correct identifications; partial compatibility awareness.	Mostly accurate identifications; minor compatibility errors.	Fully accurate and consistent use of compatibility/specifications.
System Assembly Proficiency	Incorrect sequence; misplacements; build incomplete/non-functional.	Partial sequence followed; several errors in placement/connections.	Mostly correct sequence and placement; functional build with minor issues.	Fully correct sequence and placement; complete and functional on first attempt.
Diagnostic and Troubleshooting	Fails to identify or resolve issues; random attempts.	Identifies some issues but applies incomplete or inefficient solutions.	Correctly diagnoses most issues; logical fixes with minor inefficiencies.	Accurately diagnoses and resolves all issues systematically and efficiently.
Strategic Decision-Making	Ignores constraints; poor or inappropriate decisions.	Considers some constraints but makes suboptimal decisions.	Balances most constraints effectively; solutions generally acceptable.	Consistently optimizes decisions while balancing all constraints.

Author contributions **Manuel B. Garcia:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethical standards All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Competing Interests The author is affiliated with the department that developed the MILES Virtual World platform utilized in this study. This institutional connection may be perceived as a potential source of bias. However, the study was conducted with methodological rigor, and all procedures for data collection, analysis, and interpretation were carried out independently to ensure objectivity. No financial gain, external sponsorship, or commercial interest is associated with the development or evaluation of the platform.

References

- Abbasi, S., & Kazi, H. (2020). Stealth Assessment in Serious Games to Improve OO Learning Outcomes. *2019 International Conference on Advances in the Emerging Computing Technologies (AECT)*. <https://doi.org/10.1109/AECT47998.2020.9194170>
- Adarkwah, M. A., & Islam, A. Y. M. A. (2025). Enhancing entrepreneurial skills of persons with disabilities (PWDs) through immersive learning in the metaverse: A design thinking approach. *Disability and Rehabilitation: Assistive Technology*, 20(7), 2103–2117. <https://doi.org/10.1080/17483107.2025.2524609>
- Aiello, P., D'Elia, F., Di Tore, S., & Sibilio, M. (2012). A constructivist approach to virtual reality for experiential learning. *E-Learning and Digital Media*, 9(3), 317–324. <https://doi.org/10.2304/elea.2012.9.3.317>
- Ammendola, M., Memeo, R., & Al Ansari, M. (2024). The revolution of metaverse in surgery: A mini-review with video. *Updates in Surgery*, 76(5), 1699–1703. <https://doi.org/10.1007/s13304-024-01960-x>
- Arif, Y. M., Nugroho, F., Aini, Q., Fauzan, A. C., & Garcia, M. B. (2024). A Systematic Literature Review of Serious Games for Physical Education: Technologies, Implementations, and Evaluations. In *Global Innovations in Physical Education and Health*. IGI Global. <https://doi.org/10.4018/979-8-3693-3952-7.ch001>
- Baudier, P., De Vassoigne, T., Arami, M., Delannoy, A., & Germon, R. (2025). Embracing the metaverse: User perception and acceptance of the metaverse in education. *Computers in Human Behavior*, 173, 1–13. <https://doi.org/10.1016/j.chb.2025.108804>
- Beer, P., & Mulder, R. H. (2020). The effects of technological developments on work and their implications for continuous vocational education and training: A systematic review. *Frontiers in Psychology*, 11, 1–19. <https://doi.org/10.3389/fpsyg.2020.00918>
- Bowden, J.-H., Tickle, L., & Naumann, K. (2021). The four pillars of tertiary student engagement and success: A holistic measurement approach. *Studies in Higher Education*, 46(6), 1207–1224. <https://doi.org/10.1080/03075079.2019.1672647>
- Bringula, R. P., Canseco, J. V. T., Durolofo, P. L. J., Villanueva, L. C. A., & Caraos, G. M. (2022). Effects of computer-based training in computer hardware servicing on students' academic performance. *International Journal of Technology-Enabled Student Support Services*, 12(1), 1–13. <https://doi.org/10.4018/IJTESS.317410>
- Caroux, L., & Pujol, M. (2024). Player enjoyment in video games: A systematic review and meta-analysis of the effects of game design choices. *International Journal of Human-Computer Interaction*, 40(16), 4227–4238. <https://doi.org/10.1080/10447318.2023.2210880>

- Chen, X., Zou, D., Xie, H., & Wang, F. L. (2023). Metaverse in education: Contributors, cooperations, and research themes. *IEEE Transactions on Learning Technologies*, 16(6), 1111–1129. <https://doi.org/10.1109/TLT.2023.3277952>
- Christou, G. (2014). The interplay between immersion and appeal in video games. *Computers in Human Behavior*, 32, 92–100. <https://doi.org/10.1016/j.chb.2013.11.018>
- Chu, X., Wang, M., Spector, J. M., Chen, N.-S., Chai, C. S., Hwang, G.-J., & Zhai, X. (2025). Enhancing the flipped classroom model with generative AI and metaverse technologies: Insights from lag sequential and epistemic network analysis. *Educational Technology Research and Development*, 73(3), 1539–1567. <https://doi.org/10.1007/s11423-025-10457-2>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computers & Education: X Reality*, 4, 1–11. <https://doi.org/10.1016/j.cexr.2024.100053>
- Damaševičius, R., & Sidekerskienė, T. (2024). Virtual worlds for learning in metaverse: A narrative review. *Sustainability*, 16(5), 1–41. <https://doi.org/10.3390/su16052032>
- Dobrowolski, P., Skorko, M., Pochwatko, G., Myśliwiec, M., & Grabowski, A. (2021). Immersive virtual reality and complex skill learning: Transfer effects after training in younger and older adults. *Frontiers in Virtual Reality*, 1, 1–8. <https://doi.org/10.3389/frvir.2020.604008>
- Dubinsky, J. M., & Hamid, A. A. (2024). The neuroscience of active learning and direct instruction. *Neuroscience & Biobehavioral Reviews*, 163, 1–21. <https://doi.org/10.1016/j.neubiorev.2024.105737>
- Fajaryati, N., Budiyo, Akhyar, M., & Wiranto (2020). The Employability Skills Needed To Face the Demands of Work in the Future: Systematic Literature Reviews. *Open Engineering*, 10(1), 595–603. <https://doi.org/10.1515/eng-2020-0072>
- Fang, Y., Li, T., Huynh, L., Christliff, K., Roscoe, R. D., & McNamara, D. S. (2023). Stealth literacy assessments via educational games. *Computers*, 12(7), 1–15. <https://doi.org/10.3390/computers12070130>
- Garcia, M. B. (2024). Profiling the skill mastery of introductory programming students: A cognitive diagnostic modeling approach. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-13039-6>
- Garcia, M. B. (2025). Teachers in the metaverse: The influence of avatar appearance and behavioral realism on perceptions of instructor credibility and teaching effectiveness. *Interactive Learning Environments*, 33(7), 4334–4350. <https://doi.org/10.1080/10494820.2025.2462144>
- Garcia, M. B., Adao, R. T., Pempña, E. B., Quejado, C. K., & Maranan, C. R. B. (2023). MILES Virtual World: A Three-Dimensional Avatar-Driven Metaverse-Inspired Digital School Environment for FEU Group of Schools. *Proceedings of the 7th International Conference on Education and Multimedia Technology*. <https://doi.org/10.1145/3625704.3625729>
- Garcia, M. B., Quejado, C. K., Maranan, C. R. B., Ualat, O. N., & Adao, R. T. (2024). Valentine's Day in the Metaverse: Examining School Event Celebrations in Virtual Worlds Using an Appreciative Inquiry Approach. *Proceedings of the 8th International Conference on Education and Multimedia Technology*. <https://doi.org/10.1145/3678726.3678739>
- Garg, S., Sekhar, C., & Kumar, D. (2025). Harnessing the metaverse for corporate training: Identifying critical competencies through SF-BBWM and GINA. *Journal of Management & Organization*, 31(5), 2572–2589. <https://doi.org/10.1017/jmo.2025.10019>
- Hammouda, S. B., Maoua, M., & Bouchahma, M. (2025). The effectiveness of VR-based human anatomy simulation training for undergraduate medical students. *BMC Medical Education*, 25(1), 1–14. <https://doi.org/10.1186/s12909-025-07402-5>
- Hunt, C. L., Sun, Y., Wang, S., Shehata, A. W., Hebert, J. S., Gonzalez-Fernandez, M., & Thakor, N. V. (2023). Limb loading enhances skill transfer between augmented and physical reality tasks during limb loss rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 20(1), 1–14. <https://doi.org/10.1186/s12984-023-01136-5>
- International Labour Organization (2021). *Global Framework on Core Skills for Life and Work in the 21st Century*. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@ed_emp/@emp_ent/documents/publication/wcms_813222.pdf
- Iqbal, M. Z., & Campbell, A. G. (2023). Metaverse as tech for good: Current progress and emerging opportunities. *Virtual Worlds*, 2(4), 326–342. <https://doi.org/10.3390/virtualworlds2040019>
- Jerrim, J. (2023). Test anxiety: Is it associated with performance in high-stakes examinations? *Oxford Review of Education*, 49(3), 321–341. <https://doi.org/10.1080/03054985.2022.2079616>

- Lee, S.-M., & Ahn, T. (2025). L2 learner experiences in a playful constructivist metaverse space. *ReCALL*, 37(1), 129–145. <https://doi.org/10.1017/S0958344024000235>
- Li, L. (2024). Reskilling and upskilling the future-ready workforce for Industry 4.0 and beyond. *Information Systems Frontiers*, 26(5), 1697–1712. <https://doi.org/10.1007/s10796-022-10308-y>
- Li, M., & Yu, Z. (2023). A systematic review on the metaverse-based blended English learning. *Frontiers in Psychology*, 13, 1–15. <https://doi.org/10.3389/fpsyg.2022.1087508>
- Li, Y., Chen, D., & Deng, X. (2024). The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PLoS ONE*, 19(1), 1–21. <https://doi.org/10.1371/journal.pone.0294350>
- Liu, C., Meng, S., Zheng, W., & Zhou, Z. (2025). Research on the impact of immersive virtual reality classroom on student experience and concentration. *Virtual Reality*, 29(2), 1–29. <https://doi.org/10.1007/s10055-025-01153-w>
- Miller, M., Kong, A., & Oh, J.-E. (2024). Edu-Metaverse design: Perspectives of undergraduate learners. *Computers & Education: X Reality*, 5, 1–10. <https://doi.org/10.1016/j.cexr.2024.100079>
- Min, W., Frankosky, M. H., Mott, B. W., Rowe, J. P., Smith, A., Wiebe, E., & Lester, J. C. (2020). Deep-Stealth: Game-based learning stealth assessment with deep neural networks. *IEEE Transactions on Learning Technologies*, 13(2), 312–325. <https://doi.org/10.1109/TLT.2019.2922356>
- Mislevy, R. J. (2013). Evidence-centered design for simulation-based assessment. *Military Medicine*, 178(10), 107–114. <https://doi.org/10.7205/milmed-d-13-00213>
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3–62. https://doi.org/10.1207/S15366359MEA0101_02
- Mohammed, S. Y., Aljanabi, M., & Gadekallu, T. R. (2024). Navigating the nexus: A systematic review of the symbiotic relationship between the metaverse and gaming. *International Journal of Cognitive Computing in Engineering*, 5, 88–103. <https://doi.org/10.1016/j.ijcce.2024.02.001>
- Newton, S., Alemdar, M., Rutstein, D., Edwards, D., Helms, M., Hernandez, D., & Usselman, M. (2021). Utilizing evidence-centered design to develop assessments: A high school introductory computer science course. *Frontiers in Education*, 6, 1–19. <https://doi.org/10.3389/educ.2021.695376>
- Ng, D. T. K. (2022). What is the metaverse? Definitions, technologies and the community of inquiry. *Australasian Journal of Educational Technology*, 38(4), 190–205. <https://doi.org/10.14742/ajet.7945>
- Nofal, A. B., Ali, H., Hadi, M., Ahmad, A., Qayyum, A., Johri, A., & Qadir, J. (2025). AI-enhanced interview simulation in the metaverse: Transforming professional skills training through VR and generative conversational AI. *Computers and Education: Artificial Intelligence*, 8, 1–11. <https://doi.org/10.1016/j.caeai.2024.100347>
- Onu, P., Pradhan, A., & Mbohwa, C. (2024). Potential to use metaverse for future teaching and learning. *Education and Information Technologies*, 29(7), 8893–8924. <https://doi.org/10.1007/s10639-023-12167-9>
- Peter, D., Peter, M., & Peter, P. (2025). Mapping the landscape of TVET education: A global bibliometric analysis. *Asian Education and Development Studies*. <https://doi.org/10.1108/AEDS-07-2024-0141>
- Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education*, 179, 1–16. <https://doi.org/10.1016/j.compedu.2021.104429>
- Poláková, M., Suleimanová, J. H., Madžík, P., Copuš, L., Molnárová, I., & Polednová, J. (2023). Soft skills and their importance in the labour market under the conditions of industry 5.0. *Heliyon*, 9(8), 1–20. <https://doi.org/10.1016/j.heliyon.2023.e18670>
- Popov, V., Mateju, N., Jeske, C., & Lewis, K. O. (2024). Metaverse-based simulation: A scoping review of charting medical education over the last two decades in the lens of the 'Marvelous medical education machine.' *Annals of Medicine*, 56(1), 1–26. <https://doi.org/10.1080/07853890.2024.2424450>
- Qian, Y., Wang, J., & Cai, Y. (2023). Revolutionizing educational landscapes: A systematic review of metaverse applications, paradigms and emerging technologies. *Cogent Education*, 10(2), 1–18. <https://doi.org/10.1080/23311186X.2023.2264006>
- Rahimi, S., & Shute, V. J. (2024). Stealth assessment: A theoretically grounded and psychometrically sound method to assess, support, and investigate learning in technology-rich environments. *Educational Technology Research and Development*, 72(5), 2417–2441. <https://doi.org/10.1007/s11423-023-10232-1>
- Rahimi, S., Almond, R., Ramírez-Salgado, A., Wusylko, C., Weisberg, L., Song, Y., & Wright, E. (2024). Competency model development: The backbone of successful stealth assessments. *Journal of Computer Assisted Learning*, 40(6), 2772–2789. <https://doi.org/10.1111/jcal.13025>

- Rudolph-Solero, T., Bajos-Ariza, F., Lorenzo-Álvarez, R., Domínguez-Pinos, D., Ruiz-Gómez, M. J., & Sendra-Portero, F. (2025). Problem-based learning in radiology achieves similar results in classroom and metaverse settings. *Insights into Imaging*, 16(1), 1–12. <https://doi.org/10.1186/s13244-025-01987-7>
- Ryan, G. V., Callaghan, S., Rafferty, A., Higgins, M. F., Mangina, E., & McAuliffe, F. (2022). Learning outcomes of immersive technologies in health care student education: Systematic review of the literature. *Journal of Medical Internet Research*, 24(2), 1–13. <https://doi.org/10.2196/30082>
- Santilli, T., Ceccacci, S., Mengoni, M., & Giaconci, C. (2025). Virtual vs. traditional learning in higher education: A systematic review of comparative studies. *Computers & Education*, 227, 1–31. <https://doi.org/10.1016/j.compedu.2024.105214>
- Shute, V. J., & Rahimi, S. (2021). Stealth assessment of creativity in a physics video game. *Computers in Human Behavior*, 116, 1–13. <https://doi.org/10.1016/j.chb.2020.106647>
- Shute, V. J., Ventura, M., Bauer, M., & Zapata-Rivera, D. (2009). Melding the Power of Serious Games and Embedded Assessment to Monitor and Foster Learning: Flow and Grow. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious Games: Mechanisms and Effects*. <https://doi.org/10.4324/9780203891650-31>
- Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M.-W. (2016). Advances in the science of assessment. *Educational Assessment*, 21(1), 34–59. <https://doi.org/10.1080/10627197.2015.1127752>
- Shute, V. J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, 63, 106–117. <https://doi.org/10.1016/j.chb.2016.05.047>
- Shute, V., Rahimi, S., & Emihovich, B. (2017). Assessment for Learning in Immersive Environments. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 71–87). Springer Singapore. https://doi.org/10.1007/978-981-10-5490-7_5
- Sidhu, M. S., Mousakhani, S., Lee, C. K., & Sidhu, K. K. (2024). Educational impact of metaverse learning environment for engineering mechanics dynamics. *Computer Applications in Engineering Education*, 32(5), Article e22772. <https://doi.org/10.1002/cae.22772>
- Singh, M., Sun, D., & Zheng, Z. (2024). Enhancing university students' learning performance in a metaverse-enabled immersive learning environment for STEM education: A community of inquiry approach. *Future in Educational Research*, 2(3), 288–309. <https://doi.org/10.1002/fer3.56>
- Tong, W., Xia, M., & Qu, H. (2024). Exploring Stage Lighting Education in Metaverse. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3613905.3650924>
- Tushar, H., & Sooraksa, N. (2023). Global employability skills in the 21st century workplace: A semi-systematic literature review. *Heliyon*, 9(11), 1–14. <https://doi.org/10.1016/j.heliyon.2023.e21023>
- Uddin, M., Manickam, S., Ullah, H., Obaidat, M., & Dandoush, A. (2023). Unveiling the metaverse: Exploring emerging trends, multifaceted perspectives, and future challenges. *IEEE Access*, 11, 87087–87103. <https://doi.org/10.1109/ACCESS.2023.3281303>
- UNESCO (2022). *Transforming Technical and Vocational Education and Training for Successful and Just Transitions: UNESCO Strategy 2022–2029*. <https://doi.org/10.54675/EUDU5854>
- Varghese, V. V. V., & Renumol, V. G. (2024). Video games for assessing computational thinking: A systematic literature review. *Journal of Computers in Education*, 11(3), 921–966. <https://doi.org/10.1007/s40692-023-00284-w>
- Vate-U-Lan, P., & Cahill, D. (2024). Advancing academic achievement: Integrating society 5.0 and game-based learning in metaverse space for undergraduates in Bangkok, Thailand. *Cogent Education*, 11(1), 1–15. <https://doi.org/10.1080/2331186X.2024.2385239>
- Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-017-0062-1>
- World Economic Forum (2025). *The Future of Jobs Report 2025*. <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>
- Yang, C., Zhang, J., Hu, Y., Yang, X., Chen, M., Shan, M., & Li, L. (2024). The impact of virtual reality on practical skills for students in science and engineering education: A meta-analysis. *International Journal of STEM Education*, 11(1), 1–15. <https://doi.org/10.1186/s40594-024-00487-2>
- Zhong, Z., & Juwaheer, S. (2024). Digital competence development in TVET with a competency-based whole-institution approach. *Vocation, Technology & Education*, 1(2), 1–14. <https://doi.org/10.54844/vte.2024.0591>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Manuel B. Garcia^{1,2,3} 

✉ Manuel B. Garcia
mbgarcia@feutech.edu.ph; manuelgarciaph@yahoo.com

¹ College of Education, University of the Philippines Diliman, Quezon City, Philippines

² Graduate School of Education, Korea University, Seoul, South Korea

³ Educational Innovation and Technology Hub, FEU Institute of Technology, Manila, Philippines