

# AI-Enabled MetaHuman Interviewer: A Serious Game for Student Interview Skill Development

Samiksha Lahu Gayake<sup>1</sup>, Prof. Urmila B. Shelake<sup>2</sup>, Dr. H. B. Jadhav<sup>3</sup>

<sup>1,2,3</sup>Department of Computer Engineering,  
Adsul's Technical Campus, Chas, India

**Abstract:** Student job interviews require a combination of content knowledge, communication, confidence, listening, professional behavior, and the ability to adapt to unexpected questions. Conventional preparation is often limited by access to human interviewers, repeated practice opportunities, and the difficulty of obtaining immediate, consistent feedback. This paper proposes an AI-Powered MetaHuman Interviewer, designed as a serious game in which a student participates in realistic job-interview scenarios with a photorealistic virtual interviewer. The proposed system combines an Unreal Engine-based MetaHuman, speech recognition, text-to-speech, a large language model (LLM) dialogue engine, a structured interview rubric, adaptive scenario selection, and post-session feedback. The game is designed to move beyond a static chatbot by introducing embodied social interaction, progressive difficulty, repeated practice, and measurable skill development. The paper presents the system architecture, game mechanics, question-generation strategy, response analysis pipeline, scoring model, evaluation methodology, privacy and bias safeguards, accessibility considerations, and implementation roadmap. Existing research on serious games, virtual agents, virtual-reality interview training, and AI-assisted interview simulation supports the feasibility of the approach, while also identifying risks related to automated scoring and demographic bias. The proposed evaluation uses pre-test/post-test interview performance, rubric agreement with human raters, usability, engagement, and calibration metrics. No experimental results are claimed in this paper; instead, the paper defines a reproducible evaluation protocol and expected outcomes. The proposed system aims to provide scalable, low-risk interview practice while keeping AI feedback formative rather than treating it as a substitute for human hiring decisions..

**Keywords:** Artificial Intelligence, MetaHuman, Serious Game, Job Interview Training, Virtual Interviewer, Generative AI, Student Employability, Speech Analysis.

## I. INTRODUCTION

The transition from university education to employment increasingly depends on the ability of students to demonstrate both technical competence and professional communication. A job interview is not simply a test of factual knowledge. Candidates must understand a question, construct a relevant answer, explain evidence from their experience, listen actively, communicate clearly, and respond appropriately to follow-up questions. These skills are difficult to master through reading advice alone because interview performance is situated and interactive.

Digital simulation offers a way to convert interview preparation from a one-time event into a repeatable learning activity. Earlier work on serious games and virtual agents has already explored social coaching through simulated job interviews. The TARDIS framework, for example, proposed scenario-based virtual-agent training for job interviews and emphasized practice, progression, and configurable scenarios [1]. A later conceptual framework for serious role-playing games described how virtual agents can support situated learning and experimentation with communication strategies [2].

Recent generative-AI systems create an additional opportunity. An LLM can generate follow-up questions, adapt scenarios to a student's answer, and provide individualized feedback after an interview. Recent research has



demonstrated AI-supported interview simulations in virtual environments and has also shown that bias and reliability must be examined rather than assumed [4]. These findings motivate a system in which AI is used as a coach and simulator, while explicit rubrics, human oversight, and transparent feedback constrain the model.

The central idea of this paper is to combine conversational AI with a MetaHuman-based embodied interviewer. Epic Games describes MetaHuman as a framework for creating photorealistic, fully rigged digital humans in Unreal Engine, while MetaHuman Animator supports animation driven by video and audio performance capture [11]. For interview training, an embodied interviewer can provide a more socially situated interaction than text alone, while the serious-game layer supplies goals, levels, feedback loops, and repeated practice.

The contribution of this paper is therefore a complete design for an AI-Powered MetaHuman Interviewer targeted at students. The design emphasizes four principles: realism without unnecessary complexity, measurable skills instead of opaque scores, adaptive practice instead of fixed scripts, and responsible AI instead of automated hiring judgment.

## **II. PROBLEM STATEMENT AND MOTIVATION**

Many students enter recruitment processes with limited experience in formal interviews. A student may understand a technical subject but still struggle to organize an answer, give concrete examples, maintain an appropriate speaking pace, or handle a follow-up question. Human mock interviews can address these problems, but they require staff time and scheduling. Peer practice is more scalable but can vary in quality, consistency, and realism. Video recording allows self-review but does not provide an interactive interviewer who reacts to the answer.

The problem can be framed as a learning loop: practice → observe → receive feedback → revise → practice again. Conventional preparation often breaks this loop because feedback is delayed or unavailable. A serious game can close the loop by making practice immediate and repeatable. Serious games are educational systems in which game mechanisms are used for a purpose beyond entertainment. Evidence from educational serious-game research suggests potential benefits for knowledge and skills when game design is aligned with the learning objective [6].

Interview practice has a further challenge: realism. If every question is presented in exactly the same way, students can memorize responses rather than develop adaptable communication skills. Virtual interviewers address this by supporting branching dialogue and contextual reactions. The TARDIS literature explicitly explored virtual recruiters and different interview situations [1]. More recent VR interview research with college students has reported empirical evidence that agent-based interview training can improve interview-related performance, while also noting the importance of cultural adaptation [3].

### **Research Gap**

The literature demonstrates three separate capabilities: serious-game training for social skills, virtual-agent interview simulation, and generative-AI conversational feedback. The research opportunity is to integrate these capabilities into a student-focused system that treats the interview as a complete game loop rather than a chatbot conversation. The proposed system therefore connects embodied interaction, adaptive dialogue, structured scoring, progress tracking, and responsible-AI controls.

### **Research Questions**

| RQ  | Research question                                                                                                 |
|-----|-------------------------------------------------------------------------------------------------------------------|
| RQ1 | How can a MetaHuman interviewer be integrated with generative AI to create realistic student interview scenarios? |
| RQ2 | Which interview skills can be evaluated automatically without reducing the evaluation to a single opaque score?   |
| RQ3 | How can serious-game mechanics encourage repeated practice and progressive difficulty?                            |
| RQ4 | How should the system be evaluated for learning effectiveness, usability, reliability, fairness, and              |



privacy?

### III. RELATED WORK AND COMPARATIVE ANALYSIS

Existing work provides strong evidence that interview training can benefit from simulated interaction, but the approaches differ in realism, adaptability, feedback, and evaluation. Table II summarizes the main design patterns relevant to the proposed system. The comparison is descriptive and is intended to identify design opportunities rather than rank existing studies.

TABLE II

| Approach                   | Embodiment    | Dialogue               | Feedback                     | Main relevance                              |
|----------------------------|---------------|------------------------|------------------------------|---------------------------------------------|
| Traditional mock interview | Human         | Human-led              | Human feedback               | High realism; limited scalability           |
| Video/self-practice        | None          | Fixed/self-directed    | Self-review                  | Low cost; low interactivity                 |
| Serious-game virtual agent | Virtual agent | Scenario-based         | Game feedback                | Situated social-skill practice [1,2]        |
| VR interview training      | Avatar/VR     | Interactive            | Performance-oriented         | Immersion and repeated exposure [3]         |
| LLM interview chatbot      | Text/chat     | Generative             | Automated feedback           | Flexible dialogue; bias/reliability concern |
| Proposed MetaHuman system  | MetaHuman     | Generative constrained | +Rubric + narrative feedback | Embodiment + adaptability + progression     |

#### Design Lessons from Prior Research

First, realism should serve learning. High-fidelity graphics are useful only when they improve attention, social presence, or behavioral practice. Second, adaptive dialogue is valuable because students need to respond to unexpected follow-ups rather than recite memorized scripts. Third, feedback must be specific. Statements such as “speak confidently” are less actionable than feedback linked to a measurable behavior, such as reducing excessive pauses or adding a concrete example.

Fourth, repeated practice matters. The 2025 study of an agent-based VR interview system for college students reported a training design involving multiple interviewer personalities and an empirical evaluation [3]. The result is relevant to the proposed system because interview anxiety and performance may depend on the behavior of the interviewer, not only on the question content.

Fifth, automated scoring must be treated cautiously. The 2025 metaverse interview-simulation study reported potential bias associated with candidate background variations even when the interview response was held constant [4]. This is directly relevant to student training systems because an educational score can influence confidence and perceived ability. The proposed design therefore recommends that scores be accompanied by evidence, uncertainty, and human-review options.

#### Positioning of the Proposed Work

The proposed AI-Powered MetaHuman Interviewer is positioned as an educational serious game rather than an automated recruitment system. Its objective is to create practice opportunities and formative feedback. It does not determine whether a student should be hired, and it should not be used as a proxy for an employer’s selection decision. This distinction is central to the architecture and evaluation plan.



#### IV. SYSTEM OBJECTIVES AND REQUIREMENTS

##### A. Functional Objectives

The system is designed to support a complete interview-practice cycle. A student selects a target role, experience level, and difficulty. The game initializes an interview scene and a corresponding rubric. The MetaHuman interviewer introduces the context, asks an opening question, listens to the response, and selects a follow-up when appropriate. At the end of a round, the student receives structured feedback and can repeat the scenario with increased difficulty or a new question family.

TABLE III

| ID  | Requirement           | Description                                                             |
|-----|-----------------------|-------------------------------------------------------------------------|
| FR1 | Scenario selection    | Select role, industry, difficulty, and interview style.                 |
| FR2 | Interactive interview | Conduct multi-turn question-and-answer sessions.                        |
| FR3 | Adaptive follow-up    | Generate follow-ups based on the student's answer and scenario state.   |
| FR4 | Multimodal input      | Support voice input and text fallback.                                  |
| FR5 | Feedback              | Provide evidence-based comments on content and delivery.                |
| FR6 | Progress              | Track skill trends across practice sessions.                            |
| FR7 | Accessibility         | Provide captions, transcripts, adjustable speed, and alternative input. |
| FR8 | Privacy               | Allow consent, data minimization, and deletion controls.                |

##### B. Non-Functional Requirements

Responsiveness is important because conversational latency affects social presence. The system should provide visual acknowledgement while speech is being processed, rather than leaving the MetaHuman static. Reliability requires graceful degradation: if speech recognition fails, the student should be able to type the answer. Explainability requires feedback that points to observable evidence from the response. Security requires access controls and secure storage for any recorded voice, video, transcript, or score data.

The system should also be portable across different student hardware. A high-end workstation may be able to render a photoreal MetaHuman at high quality, while a campus laboratory or student laptop may require optimized assets and lower visual fidelity. The learning objective should remain available even when graphical quality is reduced.

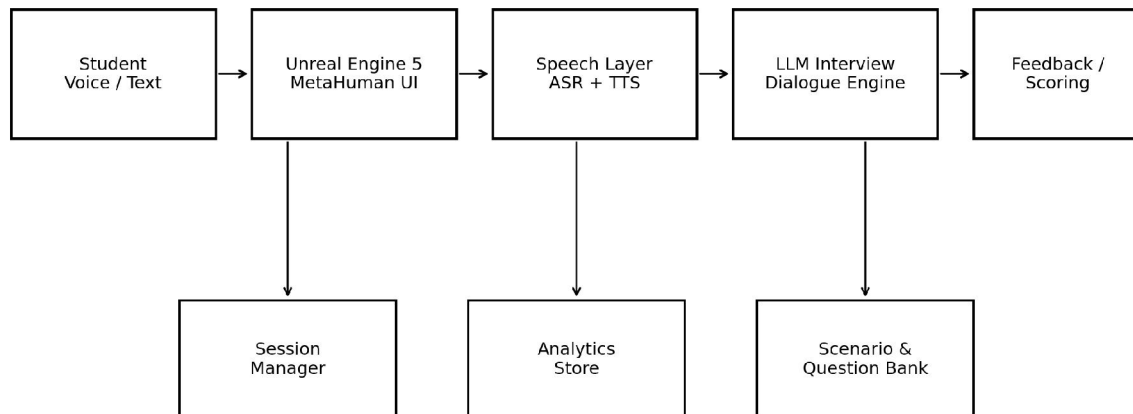
##### C. Learning Objectives

The proposed game targets six core competencies: answer relevance, answer structure, evidence/examples, clarity and concision, vocal delivery, and professional interaction. Additional role-specific competencies can be added without changing the core architecture. For example, a software-engineering scenario can include technical explanation and problem-solving, whereas a sales scenario can emphasize customer orientation and persuasion.

#### V. PROPOSED SYSTEM ARCHITECTURE

The architecture separates presentation, dialogue, analysis, and learning management. This separation is intended to make the system testable and to prevent the LLM from becoming an uncontrolled source of scores. The MetaHuman layer receives approved dialogue actions from the dialogue manager, while the scoring engine receives the transcript and structured evidence from the analysis layer.





Privacy, consent, audit logs, bias checks and human review surround the full pipeline  
Fig. 1. High-level architecture of the AI-Powered MetaHuman Interviewer.

#### A. Presentation Layer

The presentation layer is implemented in Unreal Engine and contains the interview room, camera, lighting, user interface, MetaHuman character, subtitles, timers, and game-state indicators. The MetaHuman is responsible for facial expression, gaze direction, posture, and speech synchronization. These behaviors should be driven by state variables rather than generated freely by the language model.

#### B. Conversation Layer

The conversation layer maintains interview state: role, scenario, question category, previous answers, difficulty, time budget, and allowed follow-up types. The LLM receives only the information needed to produce the next dialogue act. A constrained output schema can require the model to return a question, intent, difficulty level, and evaluation target instead of arbitrary application logic.

#### C. Analysis and Scoring Layer

The analysis layer processes the student's transcript and optional audio features. Content analysis can identify relevance, completeness, evidence, structure, and unsupported claims. Delivery analysis can measure non-clinical features such as speaking rate, pause frequency, volume consistency, and filler-word frequency. These measures should be presented as communication indicators, not as psychological or personality diagnoses.

#### D. Data Layer

The data layer stores only the information needed for progress tracking. A privacy-preserving implementation can retain derived metrics and feedback while deleting raw audio after a defined period. If recordings are required for self-review, retention should be explicit and configurable.

### VI. INTERVIEW SKILL SCORING MODEL

A central design challenge is converting a complex interview response into useful feedback without pretending that communication can be measured perfectly. The proposed solution is a multi-dimensional rubric. The overall score is optional; the primary feedback is dimension-level evidence.



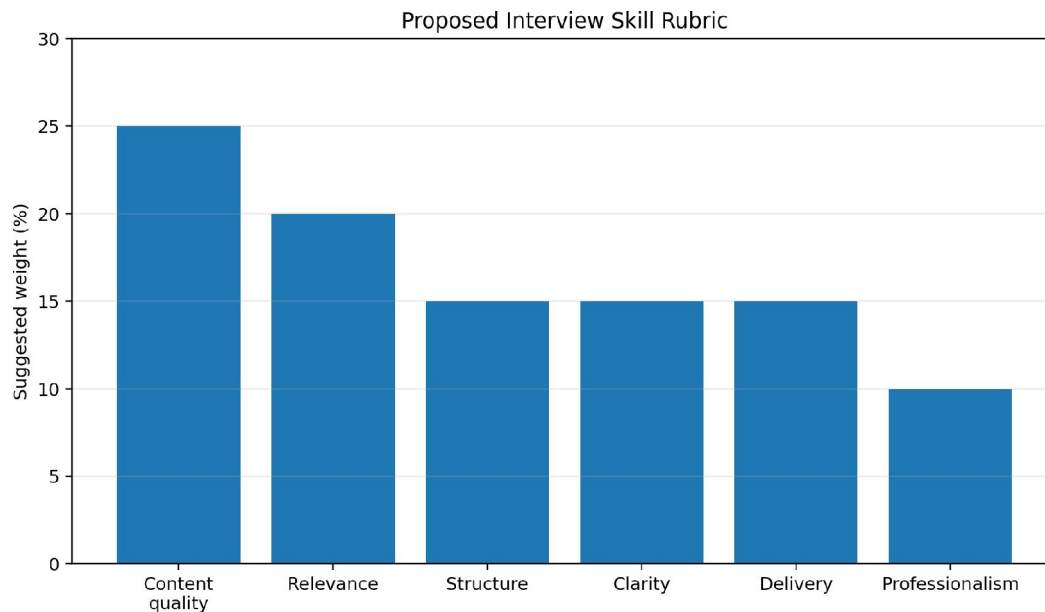


Fig. 4. Proposed weighting of interview-skill dimensions. Weights are design parameters for evaluation and should be calibrated with human raters.

#### A. Content Quality — 25%

Content quality measures whether the answer contains an understandable claim, appropriate detail, and evidence. For behavioral questions, a structured example can be evaluated for situation, action, and outcome. For technical questions, the system can evaluate conceptual correctness against a reference rubric rather than relying on surface similarity.

#### B. Relevance — 20%

Relevance measures whether the response answers the actual question. The analysis should distinguish a concise answer from a short but incomplete answer. LLM-based semantic analysis can propose a relevance assessment, but benchmark questions and human-rater validation are needed before the metric is used for learning analytics.

#### C. Structure — 15%

Structure measures logical organization and sequencing. Feedback can identify whether the answer begins with a clear point, provides supporting evidence, and ends with an outcome or conclusion.

#### D. Clarity — 15%

Clarity concerns understandable language, sentence construction, unnecessary repetition, and excessive jargon. The system should not penalize a student's regional accent or multilingual background merely because automatic speech recognition is less accurate.

#### E. Delivery — 15%

## VII. IMPLEMENTATION METHODOLOGY

The proposed prototype can be developed in iterative phases. Each phase produces a testable artifact so that technical problems can be identified before the full system is assembled.

TABLE V

| Phase               | Implementation work                       | Validation                   |
|---------------------|-------------------------------------------|------------------------------|
| 1. Scenario design  | Job roles, competency map, question bank  | Expert review                |
| 2. MetaHuman scene  | Character, room, camera, animation states | Usability walkthrough        |
| 3. Speech interface | ASR, TTS, captions, interruption handling | Word-error and latency tests |



|                    |                                                |                            |
|--------------------|------------------------------------------------|----------------------------|
| 4. Dialogue engine | LLM prompts, state machine, structured outputs | Scenario consistency tests |
| 5. Scoring         | Rubric mapping and evidence extraction         | Human-rater agreement      |
| 6. Game layer      | Levels, rewards, replay, dashboard             | Pilot user test            |
| 7. Responsible AI  | Consent, logging, bias audit, deletion         | Risk and privacy review    |

#### **A. Unreal Engine and MetaHuman**

The implementation uses Unreal Engine as the runtime environment and a MetaHuman character as the interviewer. Epic documentation describes current MetaHuman workflows for creating and integrating characters and provides animation pathways using MetaHuman Animator and related tools [11]. The prototype should begin with a small number of animation states rather than attempting fully autonomous facial behavior.

#### **B. Backend Services**

A backend service manages sessions, scenario state, prompts, model requests, transcripts, scoring outputs, and progress records. API boundaries should be explicit. The Unreal client should not contain API keys or sensitive model credentials. Authentication and authorization should be implemented on the server side.

#### **C. Data Model**

A session record can contain: anonymous student identifier, scenario identifier, rubric version, model version, question sequence, transcript, derived metrics, feedback, and timestamps. Raw audio should be optional. If stored, it should have a defined retention period. The system should separate personally identifying information from learning records whenever possible.

#### **D. Error Handling**

Speech recognition errors, network failures, LLM timeouts, and invalid model outputs should be treated as normal operating conditions. The game can fall back to text input, a predefined question, or a local scripted response. A failed AI service should not terminate a student's practice session.

### **VIII. EVALUATION METHODOLOGY**

Because this paper proposes a system rather than reporting a completed experiment, the evaluation protocol is specified in advance. The purpose is to determine whether the system improves interview performance, whether its automated assessments agree with human raters, and whether students find the experience usable and engaging.

#### **A. Participants**

A pilot study can recruit undergraduate students preparing for internships or campus recruitment. A larger study should include students from multiple disciplines and academic years. Participation should be voluntary, with informed consent and the ability to withdraw. Demographic information should be collected only when needed for fairness analysis and should not be used to personalize interview scoring.

#### **B. Study Design**

A controlled pre-test/post-test design is recommended. Participants complete a baseline interview assessed by independent human raters. The intervention group then completes several practice sessions with the MetaHuman system, while a comparison group receives the institution's usual interview-preparation material or mock-interview practice. A final interview uses new but equivalent questions. The primary outcome is change in rubric-based interview performance.

#### **C. Measures**

**TABLE VI**

| Measure               | Definition                  | Collection              |
|-----------------------|-----------------------------|-------------------------|
| Interview performance | Human-rated rubric score    | Pre/post interviews     |
| Automated agreement   | Agreement with human raters | Same recorded responses |



|                      |                                             |               |
|----------------------|---------------------------------------------|---------------|
| Usability            | Ease of use and learnability                | Questionnaire |
| Engagement           | Completion, retries, session duration       | System logs   |
| Perceived usefulness | Student perception of training value        | Questionnaire |
| Latency              | Time from response end to interviewer reply | System logs   |
| Fairness             | Score consistency across controlled groups  | Audit dataset |

#### **D. Statistical Analysis**

For pre/post comparisons, paired tests can be used within groups and between-group models can be used when a control condition is available. Effect sizes and confidence intervals should be reported in addition to p-values. Inter-rater reliability can be estimated using an appropriate agreement statistic based on the rubric scale and number of raters. Automated scoring should be compared to human ratings at both the total and dimension levels.

A key methodological requirement is to keep training and evaluation questions separate. If the LLM has seen the exact evaluation question during prompt design, performance can be inflated by memorization. The final test should therefore contain unseen but equivalent scenarios.

### **IX. PROPOSED RESULTS FRAMEWORK AND EXPECTED OUTCOMES**

No experimental results are claimed in this paper. Instead, this section defines what successful evaluation would look like and how evidence should be interpreted. The distinction is important because a system-design paper should not present anticipated improvements as measured findings.

#### **A. Learning Effectiveness**

The primary expected outcome is improvement in human-rated interview performance from pre-test to post-test. The strongest evidence would be a statistically and practically meaningful improvement in the intervention group relative to a comparison condition using unseen questions. A smaller but consistent improvement in a specific skill, such as answer structure or follow-up handling, may also be educationally meaningful.

#### **B. Automated Assessment Quality**

The system should show reasonable agreement with trained human raters. Agreement should be examined separately for content and delivery because different modalities have different reliability. For example, relevance may be easier to validate than professionalism, while speaking rate may be measured more consistently than “confidence.”

#### **C. Engagement**

A serious game should encourage repeated practice. Useful engagement indicators include the number of completed sessions, number of voluntary retries, and proportion of students returning for a second session. High engagement alone does not prove learning; it should be interpreted together with performance data.

#### **D. Usability**

Students should be able to begin an interview without extensive technical instructions. Usability evaluation should identify problems with microphone setup, speech latency, avatar behavior, navigation, feedback comprehension, and difficulty selection. Qualitative interviews can reveal issues that numerical questionnaires miss.

#### **E. Fairness and Reliability**

The system should demonstrate that equivalent answers receive broadly consistent feedback across controlled variations. Reliability testing should cover paraphrased answers, speech-recognition noise, different speaking rates, and changes in non-essential candidate descriptors. The recent literature shows that robustness and bias are central concerns for AI interview systems [4].

#### **F. Expected System-Level Benefits**

If the evaluation criteria are satisfied, the system could provide three practical benefits: scalable access to repeated interview practice, immediate structured feedback, and a safe environment for experimentation. These are expected benefits derived from the design and prior literature, not findings established by this paper.



A useful deployment dashboard would therefore report learning change, human-rater agreement, system reliability, usability, and fairness separately. Combining all dimensions into one “AI interview score” would conceal important trade-offs.

## **X. DISCUSSION**

The proposed system demonstrates how generative AI and embodied digital humans can be combined without allowing either technology to dominate the educational design. The MetaHuman creates a social interface, but the learning objective is defined by the rubric. The LLM creates conversational variation, but the scenario manager limits its scope. The game creates motivation, but progress is based on practice behaviors and skill trends rather than superficial rewards.

### **A. Why Embodiment Matters**

Text chat can be useful for practicing answer content, but an interview is inherently social. The candidate must speak to another person, manage turn-taking, listen, and respond under time pressure. A MetaHuman provides a visual target for attention and a controllable set of interviewer behaviors. Prior virtual-agent and VR studies support the broader idea that simulated social interaction can be used for communication training [1–3].

### **B. Why the System Should Not Fully Automate Judgment**

Interview performance is context-dependent. An answer may be appropriate for one organization and less appropriate for another. Automated scoring can also be affected by transcription errors, model variation, and demographic bias. Therefore, the proposed system uses AI as a formative coach. Human raters remain necessary for validating the rubric and auditing system behavior.

### **C. Cultural and Institutional Adaptation**

Interview norms vary by country, sector, organization, and role. A generic global interviewer may provide feedback that is technically coherent but culturally inappropriate. The scenario configuration should therefore allow institutions to define competency descriptions, question families, acceptable communication conventions, and role-specific expectations. This is consistent with recent research emphasizing cultural adaptation in virtual interview training [3].

### **D. Scalability**

The system can support many students because the AI interviewer can conduct sessions asynchronously. However, scalability introduces governance requirements. Model changes can alter feedback behavior across the student population. Versioning prompts, rubrics, models, and scenarios is therefore essential for reproducibility.

### **E. Relationship to Existing Evidence**

Research on serious games suggests that well-designed game-based learning can improve skills in some contexts [6]. Virtual interview systems have demonstrated feasibility and, in some studies, positive training effects [3]. AI interview simulations are increasingly capable of adaptive dialogue but also reveal bias and reliability concerns [4]. The proposed architecture combines these lessons and makes evaluation of the combined system an explicit research objective.

### **F. Practical Deployment in Universities**

A university can deploy the system through a placement-training laboratory, a controlled desktop application, or a networked lab. The first release should use a small set of validated job roles and a conservative rubric. More advanced capabilities such as facial-expression analysis can be added only after the core system demonstrates value, reliability, and user acceptance.

## **XI. LIMITATIONS AND FUTURE WORK**

The proposed design has several limitations. First, the quality of feedback depends on the quality of the rubric, prompts, speech recognition, and model. A sophisticated MetaHuman cannot compensate for weak educational criteria. Second, automated delivery metrics may be noisy and should not be interpreted as direct measures of confidence or personality. Third, photorealistic rendering increases hardware requirements and development complexity.



Fourth, LLM behavior can change between model versions. A prompt that produces stable feedback in one version may behave differently after an update. Model and prompt versioning is therefore required. Fifth, a serious game may create engagement without producing durable transfer to real interviews. The strongest future evaluation should test transfer through independent human-rated interviews, ideally with realistic job contexts.

Sixth, the proposed system does not solve the broader problem of interview fairness. It can help students practice, but it cannot guarantee that real employers use fair hiring procedures. The system should therefore avoid claims that its score predicts employment outcomes unless such predictive validity is independently demonstrated.

### **Future Work**

Future research can explore multi-interviewer panels, role-specific technical interviews, multilingual interview practice, culturally configurable scenarios, adaptive difficulty using validated learning models, and instructor dashboards. Another direction is the use of lightweight on-device speech processing to reduce cloud data transfer. Research can also compare a photorealistic MetaHuman with a simpler avatar to determine whether visual realism produces measurable learning benefits or mainly increases development cost.

A further research direction is calibration. The system could estimate its own uncertainty and abstain from scoring when the transcript is unreliable or the question is outside the rubric. This would move the system from forced judgment toward uncertainty-aware educational feedback. NIST's AI RMF emphasizes lifecycle testing, human-AI configuration, and ongoing monitoring, which align with this direction [9,10].

## **XII. CONCLUSION**

This paper presented a research and implementation framework for an AI-Powered MetaHuman Interviewer designed as a serious game for student job-interview skills. The proposed system combines a MetaHuman-based interactive environment, speech input and output, generative AI dialogue, a constrained interview state machine, multidimensional scoring, adaptive feedback, and repeated game-based practice. The design is motivated by prior research on serious games, virtual agents, VR interview training, and AI-assisted interview simulation.

The key design principle is to use AI to create practice and feedback rather than to make employment decisions. A structured rubric, evidence-based feedback, human validation, privacy safeguards, accessibility, and bias auditing are therefore integral components of the system. The paper does not claim experimental results; instead, it defines a reproducible evaluation methodology that can establish whether the prototype improves interview performance and whether its automated feedback is reliable and fair.

With careful validation, the proposed system can serve as a scalable complement to university placement training. Its main research value lies not only in the realism of the MetaHuman, but in the integration of embodied interaction, adaptive conversation, measurable learning objectives, and responsible AI into a repeatable practice loop.

## **REFERENCES**

- [1] K. Anderson, E. André, T. Baur, S. Bernardini, M. Chollet, E. Chrysafidou, I. Damian, C. Ennis, A. Egges, P. Gebhard, H. Jones, M. Ochs, C. Pelachaud, K. Porayska-Pomsta, P. Rizzo, and N. Sabouret, "The TARDIS Framework: Intelligent Virtual Agents for Social Coaching in Job Interviews," in *Advances in Computer Entertainment*, 2013, pp. 476–491. doi: 10.1007/978-3-319-03161-3\_35.
- [2] O. Othlinghaus-Wulhorst and H. U. Hoppe, "A Technical and Conceptual Framework for Serious Role-Playing Games in the Area of Social Skill Training," *Frontiers in Computer Science*, vol. 2, 2020. doi: 10.3389/fcomp.2020.00028.
- [3] W. Geng, C. Zhou, and Y. Bian, "Change gently: an agent-based virtual interview training for college students with great shyness," *Virtual Reality*, vol. 29, art. 12, 2025. doi: 10.1007/s10055-024-01076-y.



- [4] "AI-enhanced interview simulation in the metaverse: Transforming professional skills training through VR and generative conversational AI," *Computers and Education: Artificial Intelligence*, vol. 8, 2025, art. 100347. doi: 10.1016/j.caeai.2024.100347.
- [5] "Changes in Emotional and Vocal Expression in Job Interview Simulations with an AI-Enhanced Chatbot for University Students," *PubMed Central*, 2026. One-group pretest–posttest study of repeated chatbot-based interview practice with facial and vocal indicators.
- [6] L. Wang, Q. Zhao, L. Dong, H. Zhao, L. Qin, T. Deng, H. Huang, M. Li, X. Wu, and J. Liu, "The effectiveness of serious games on undergraduate nursing students' knowledge and skills: A systematic review and meta-analysis," *Nurse Education in Practice*, vol. 80, 2024, art. 104102. doi: 10.1016/j.nepr.2024.104102.
- [7] "An Affective Model for a Virtual Recruiter in a Job Interview Context," *Procedia Computer Science*, vol. 15, 2012, pp. 312–313. doi: 10.1016/j.procs.2012.10.090.
- [8] A. Bijl, B. P. Veldkamp, S. Wools, and S. de Klerk, "Serious games in high-stakes assessment contexts: a systematic literature review into the game design principles for valid game-based performance assessment," *Educational Technology Research and Development*, vol. 72, pp. 2041–2064, 2024. doi: 10.1007/s11423-024-10362-0.
- [9] E. Tabassi, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," *National Institute of Standards and Technology, NIST AI 100-1*, 2023. doi: 10.6028/NIST.AI.100-1.
- [10] C. Autio, R. Schwartz, J. Dunietz, S. Jain, M. Stanley, E. Tabassi, P. Hall, and K. Roberts, "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile," *NIST AI 600-1*, 2024. doi: 10.6028/NIST.AI.600-1.
- [11] Epic Games, "MetaHuman Documentation," *Epic Developer Community*, documentation for MetaHuman creation, animation, and Unreal Engine integration, accessed 2026.
- [12] N. Gavish and I. Izraellov, "Directive versus non-directive questions in an LLM-based job interview training," *Results in Engineering*, vol. 32, 2026, art. 112857. doi: 10.1016/j.rineng.2026.112857.

#### ACKNOWLEDGMENT

The authors acknowledge the use of the IJARSCT formatting template supplied for this manuscript and the publicly available research and technical documentation cited in the reference list. The proposed system and evaluation protocol in this paper are presented as a research design; experimental validation remains future work.

