2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #52023

Design and Development of an Adaptive Guidance Tool in Systems Engineering to Improve Stakeholder Engagement

Bryant Baldwin, University of South Alabama

Bryant Baldwin is a Systems Engineering PhD Student at the University of South Alabama in Mobile, AL. He received his B.S. and M.S. in Mechanical Engineering in 2020 and 2021.

Prof. Sean Walker

Dr. Walker is an Assistant Professor at the University of South Alabama in Mobile, AL. He received his Ph.D. in Systems Design Engineering from the University of Waterloo, in Waterloo, Canada, in 2012.

Grayson Patrick Dennis, University of South Alabama
Seethambal S Mani
Kari J Lippert, University of South Alabama

Dr. Kari Lippert, D.Sc., has over 45 years' experience as a Systems Engineer serving in various roles in a wide variety of both commercial and government positions. She is currently an Assistant Professor in Systems Engineering at the University of South Alabama. She is a non-typical systems engineer having started her academic career in the biological sciences. She then moved into theoretical chemistry and biochemical simulation, then big data and databases, then systems, then cyberspace defense and network security. Analysis, design, implementation, integration, testing, requirements management, change management, risk, architecture, and process improvement – all the aspects of systems engineering in a wide variety of situations and working with a large variety of systems. Her current research interests include Artificial Intelligence, Digital Ecosystems, Digital Twins and Virtual Humans.

WIP: Design and Development of an Adaptive Guidance Tool in Systems Engineering to Improve Stakeholder Engagement

Abstract

The increasing complexity of modern Systems Engineering requires emerging practitioners to develop systems thinking, effective communication, and adaptive reasoning. Traditional approaches to teaching Systems Engineering often lack the customized guidance needed to mature core competencies. This paper presents the work-in-progress toward the development of an Adaptive Systems Engineering Guidance Tool (ASEGT), an intelligent tutoring system equipped with the capability of providing personalized feedback for all users in all domains of Systems Engineering. ASGET applies core principles of systems design within the Concept Design phase in the product development lifecycle. Utilizing the Generalized Intelligent Framework for Tutoring (GIFT) platform, enhanced with Artificial Intelligence-driven adaptive modules, ASGET dynamically evaluates the learner's pedagogical preference and proficiency-level, then provides a tailored, individualized development methodology for the user. For this use case, learners will focus on the skills necessary to elicit requirements from the stakeholder and conduct stakeholder analyses. Using this engaging platform, users will simulate a dialogue with the stakeholder and receive assessment results based on INCOSE Competency Framework and NASA Systems Engineering Handbook. ASGET will provide recommended strategies to improve the technical maturity of the user. Expected users of this platform are students and industry professionals interested in improving their abilities. Current efforts focus on the system design, algorithmic framework, and interface development within the GIFT environment. Preliminary user evaluations are planned to assess usability, adaptability, and cognitive engagement experience. The expected result will be an AI-enhanced educational tool capable of enhancing the expertise of individuals by providing targeted mentorship in academic and professional settings.

Introduction

Modern systems engineering has been characterized by the development of complex systems, often operating under strict budget constraints, and under schedule pressures. As systems evolve from isolated components to interconnected systems of systems (SoS) ranging from simple functionality of a rotary telephone to multifaceted connectivity of smart eyewear, the demands placed on emerging practitioners escalate. The pace of technological advancements will continue to accelerate with no foreseeable relief. Simultaneously, growing market demands, and global resource constraints have rendered a new landscape where efficiency is paramount. Consequently, the margin of error in successful system development and deployment is shrinking.

Effective systems engineers demonstrate high-level systems thinking, effective communication, synergy, and adaptability [1-3]. However, due to the emphasis of technical competencies, traditional educational approaches to systems engineering (SE) often fail to keep pace with the needs of industry [3-6]. Standard approaches frequently rely on (1) a static, "onesize-fits-all" educational methodology to provide guidance for emerging practitioners or (2) ad-hoc, on-the-job mentorship to mature individual core competencies [7]. However, expert mentorship is not always available, knowledge transfer programs, generalized educational training, and institutional instruction may fail to reach the specific needs of the learner. There is a critical need for educational tools that can bridge the gaps by offering personalized, scalable guidance.

Background

Intelligent Tutoring Systems (ITS) have garnered attention across academic, commercial, and industrial settings. ITSs are computer-based systems designed to provide adaptive instruction and feedback to learners. These systems leverage artificial intelligence, cognitive modeling, and data-driven analytics to: (a) assess the learner's knowledge and learning style, (b) adapt instruction based on acquired analytics, (c) provide tailored feedback, (d) track progress and learning outcome [8, 9]. To date, ITSs have been applied to a wide range of domains that include mathematics, medical, language science, finance, and military [8, 10-12] but seldom in the domain of Systems Engineering. Table 1 provides the typical breakdown of these systems [8, 13]:

Table 1: ITS Breakdown

Component	Definition
Expert Module	Contains Knowledge of the subject being taught (domain knowledge)
Student Model Module	Built to maintain the learner's profile (level of knowledge, behaviors, learning styles, knowledge deficiency, etc.)
Tutoring Module	Detects the knowledge deficiency in students and focuses on the strategies and methods of teaching for compensating for the identified shortage. This is also known as the Pedagogical Model.
User Interface	The forward-facing communication part of the ITS for controlling the interaction between the user and system.

The effectiveness of ITS in improving student outcomes has been well-documented over the past several decades. SCHOLAR is an ITS by Jamie Carbonell and considered the earliest capability demonstration [8, 9, 13]. This program was developed to evaluate the learner's knowledge of South American geography. It displayed the functionality of inquiring, detecting misspellings, awarding partial credit, and tracking context [14]. Koedinger et al. reported the large-scale implementation and success of PUMP Algebra Tutor (PAT), designed to support mathematical modeling in urban high schools in Pittsburgh, Philadelphia [10]. PAT emphasizes situated learning approach, guiding students through complex, real-world problem-solving scenarios. The researchers found that students utilizing the ITS outperformed their peers in traditional classrooms across all metrics, demonstrating 15-25% higher scores on standardized mathematics tests and 100% increase in performance on authentic problem-solving tests. In a comprehensive review of 50 controlled evaluations, researchers assessed the efficacy of ITS implementations against conventional instruction. Results indicate that ITSs can significantly improve learning outcomes, bridging the gap between mass education and personalized mentorship [15]. Their meta-analysis yielded compelling results of 0.66 standard deviation improvement on test score over conventional levels, or from 50^{th} to the 75^{th} percentile. While human tutoring typically raises student test scores by about 0.40 standard deviations over control level [16], it is understood that Digital Tutoring systems when correctly implemented will improve educational outcomes more significantly [15].

In the industrial domain, the SHERLOCK system, developed by Alan Lesgold and Sherries Gott, investigated the potential of ITS technology to enhance complex skill acquisition of diagnostics and troubleshooting of avionics technicians. SHERLOCK demonstrated remarkable results in which senior apprentices (technicians of three years of on-the-job experience) outperformed a group of senior technicians having 10 or more years of experience on a controlled evaluation. Studies concluded that approximately 25 hours of SHERLOCK training was equivalent to four years of on-the job training [12, 17]. Lastly, DARPA's Digital Tutor is another tool that is considered a great advancement in this area. Thirty-two students completed a 16-week program of instruction integrated with 2 weeks of a prototype ITS developed by DARPA, and upon completion, these trainees were given a series of simulated trouble-shooting problems of varying difficulty. When compared to technicians with 12 years of experience, the leaners significantly outperformed the veterans [18]. This result highlights the potential for well-designed ITS to accelerate the acquisition of expertise.

The primary objective of this work is to examine how voluntary engagement with such ITS tools influences the development of core systems engineering competencies and enhances product development outcomes. This paper presents ongoing work on the creation of such a tool, with a particular emphasis on a SoS tool capable of providing a scenario-informed, guidance and competency evaluation.

Methodology

The methodology integrates iterative prototyping within the Generalized Intelligent Framework for Tutoring (GIFT) environment, formative evaluation with representative users, and quantitative/qualitative analyses of learning outcomes and user experience. GIFT is built around three primary functions [19]:

- Authoring: Provides the tools and methods to make building computer-based tutoring systems affordable and accessible. The tools include domain-specific knowledge configuration tools and instructional strategy development tools.
- Instructional Management: Provides a modular pedagogical engine that delivers instructional actions, learning theory, and motivation theory.
- Analysis Construct: A data-capture and evaluation service that allows researchers to assess the effectiveness of the tutoring and learning outcomes.

This paper discusses the design science and user-centered approach to develop and evaluate the effectiveness of the Adaptive Systems Engineering Guidance Tool (ASEGT). The overarching goal of this work is to: (1) optimize system engineering competencies in the Concept Development phase, (2) deliver adaptive, individualized guidance, and (3) assess usability, adaptability, and cognitive engagement in realistic stakeholder-elicitation tasks. The system includes integrated modules for evaluating learning styles, core competencies, and design package evaluation. The user profile is vital as this tool adapts as it collects data on engagement metrics, skill mastery assessments, and learning path progression. The platform delivers personalized content and continuously adjusts learning paths based on user feedback.

This observational experimental study investigates the adoption and effectiveness of the ASEGT. This study will align with established methodologies of ITS system development, deployment, and evaluation. This section outlines ASEGT's interfaces, the simulation module, expected participants, and results analysis. The goal is to develop a comprehensive understanding of how users interact with the simulation tool, the impact of adaptive learning on their systems engineering competencies, and the overall effectiveness of the guidance provided. By analyzing user feedback and user performance metrics (e.g., learning outcomes metrics, user performance metrics, and user engagement metrics), the study seeks to identify best practices for enhancing the tool's functionality.

Implementation within GIFT

An event pipeline was designed and implemented within the GIFT platform to manage the flow of instructional and system events across multiple modules. This pipeline orchestrates interactions between modules ensuring seamless data exchange and real-time adaptability. Figure 1 provides an illustration of the event pipeline.

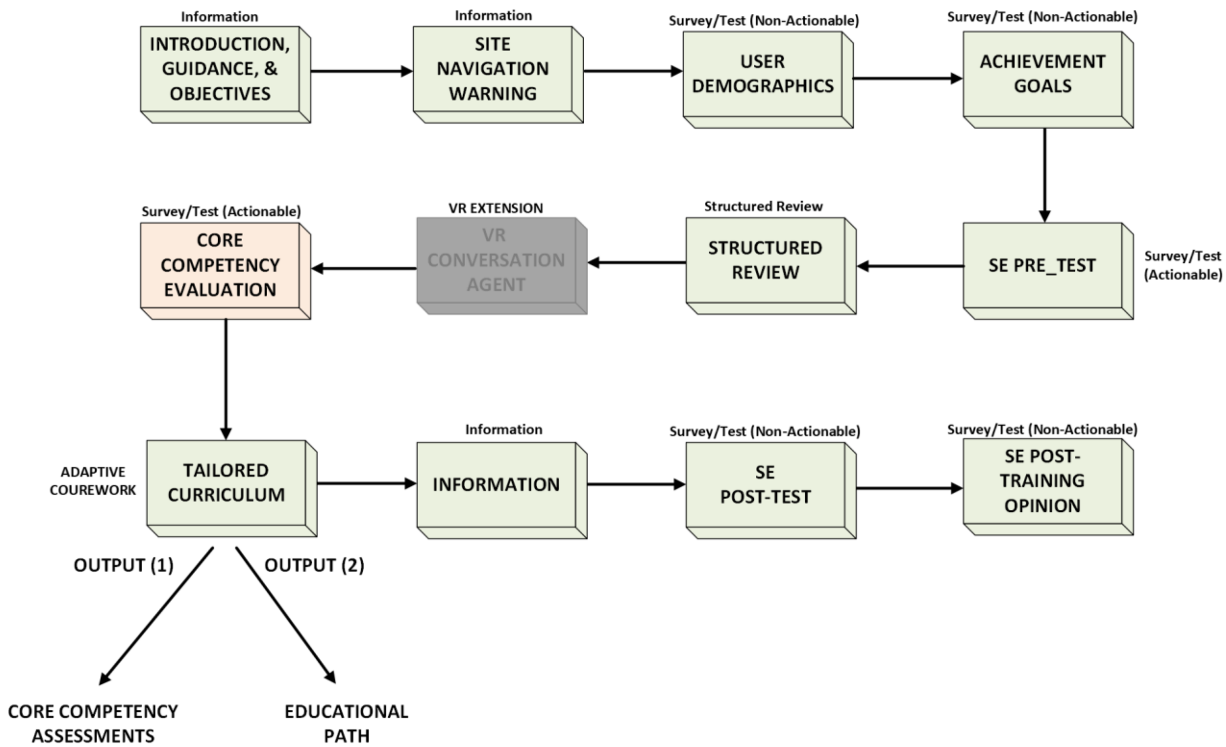

Figure 1 ASEGT GIFT System Design Architecture

As illustrated in Figure 1, the current iteration of the system features 11 active modules. These modules include information blocks, actionable and nonactionable surveys, structured reviews, and adaptive coursework. An additional module (VR Extension) is currently under development and will be discussed in a future paper. The system progresses from gathering demographics and experience level to providing a tailored curriculum to strengthen core competencies.

The Stakeholder Simulation Module is the centerpiece of this study. This tool is integrated into a broader ASEGT which is designed with a sequential pipeline that guides users through a comprehensive SE learning experience. The journey begins with a foundational introduction and site navigation warning, ensuring users are oriented to the platform's objectives and interface guidelines. The introductory page provides an overview of the platform and its objectives. While the Site Navigation Warning page provides the user with critical information regarding how to navigate the web browser. To tailor the experience, the system will capture user demographics to serve as a reference point for the user and establish a baseline for the system. This module collects baseline information such as educational background and professional experience. Following the Demographic module is the Achievement Goals module which enables users to specify their learning objectives, preferred learning style, and desired outcomes. At present, this information is stored for reference but is not actively utilized to influence the system's adaptive behavior.

The core pedagogical cycle starts with a SE Pre-Test/Assessment module which establishes a technical baseline for each participant. By assessing initial technical competency, the system creates a personalized starting point, storing results to categorize users into distinct learning journeys at a later stage. Following the assessment, users enter a Structured Review, where they can evaluate their performance and address gaps in prior knowledge. If the user response is incorrect, the system triggers a Remediation module, delivering targeted media (article or video) specifically related to the lesson.

Upon finishing the Remediation phase, the learner will experience the Stakeholder Simulation Tool (the focus of this publication) which utilizes scenario-based assessments aligned with INCOSE standards to measure core Systems Engineering competencies. The Graphical User Interface (GUI), as shown in Figure 2, facilitates a text-based simulation of dynamic stakeholder interaction. For this specific lesson, users are focusing on requirements definition and negotiation skills; therefore, the simulation requires learners to navigate conflicting stakeholder priorities to reach a technically sound consensus. Furthermore, currently the system is currently being expanded to include a module that steers users toward either text-based simulation or Virtual Reality (VR) Conversation Agent. While both environments provide the same stakeholder simulation, The VR agent introduces an additional challenge by requiring users to vocalize their responses rather than typing them.

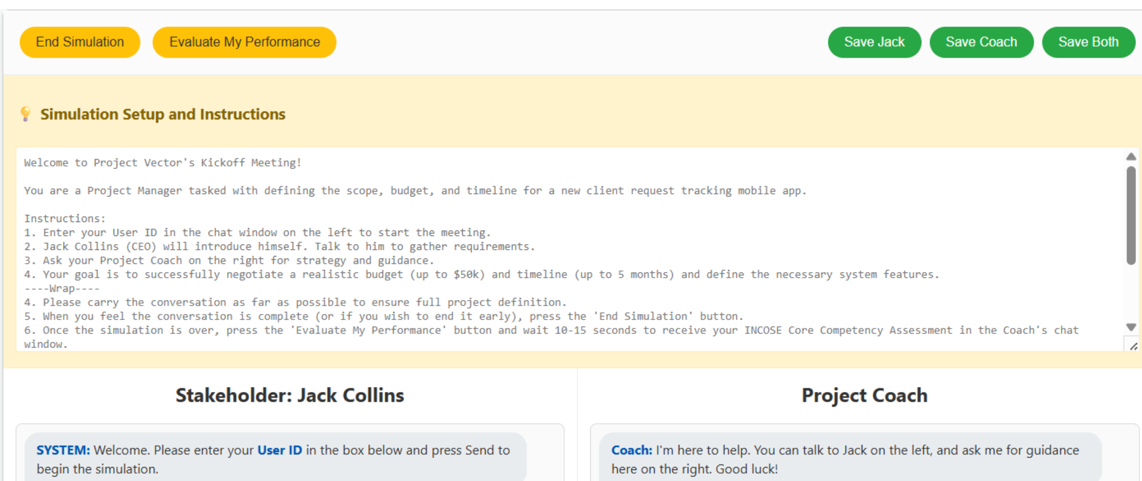

Figure 2 Screenshot of Landing Page of Simulation Tool

Based on results of the pre-assessment and simulation, the Tailored Curriculum module guides users through specific recommended paths. This module delivers adaptive content and exercises based on user performance and learning preferences identified in earlier modules. Primary outputs of this module are documentation for (1) core competency assessments, and (2) educational pathway recommendations (these features are currently under development). At the end, the user completes the experience through a series of modules: an Informational Block that provides future plans for the tool and contact information, an SE-Post Assessment to evaluate knowledge gain, and an SE-Post Training Opinion survey to capture usability feedback.

Figure 2 shows the initial interface of the stakeholder simulation tool. The left dialogue panel enables the user to interact with the simulated stakeholder while the right panel provides access to the project coach for real-time guidance. These panels represent the functional implementation of Dialogue Emulator and the Conversational Assistor subsystems. The Stakeholder is programmed to only engage the user and cannot view the dialogue from the Project Coach; however, the project coach is designed to view the ensure dialogue between the user and the stakeholder. Added features to this tool are: "Save Jack Log, "Save Coach Log," and "Save Both Logs." This added functionality will allow learners to review their performance post-engagement. Upon exit of the platform, the tool will wipe all scenario history. Outputting the logs will also allow for the integration into the overarching ITS.

To improve system accuracy and effectively assess the skill level of the user, it is recognized that traditional knowledge evaluation is limited and often insufficient for SE Competencies (e.g., Requirements Definition, Negotiation, Systems Thinking, Emotional Intelligence). The INCOSE Systems Engineering Competency Framework distinguishes between knowing a fact and having the capability to use it. This is reflected in the proficiency levels (Awareness, Supervised Practitioner, Practitioner, Lead Practitioner, and Expert) [2]. A knowledge evaluation can effectively assess conceptual understanding and confirm awareness-level proficiency; however, higher-order competencies like Negotiation and Systems Thinking require advanced proficiency that can only be assessed through demonstrated performance or behavioral observation. This distinction is validated in literature that discuss the competency gap where engineers "possess a vocabulary of semantics but lack the skills to apply what they learned" [20]. Furthermore, research promotes the need to transition the education role to the "guide on the side" to effectively scaffold learning and enable practitioners to develop independent critical thinking skills [4].

In particular, the immediate focus of the immediate study is to evaluate the Stakeholder Simulation Module. This tool is designed with a similar architecture of a traditional ITS system with the capabilities of: (1) Manage Access, (2) Manage User Engagement, (3) Perform simulation, (4) Perform evaluation, (5) Manage Session. The focus of this paper is to describe the pedagogical approach and evaluation strategy for this study.

Data Analysis

The study utilizes three primary metrics: Net Promoter Score (NPS), Mean Scores with Standard Deviation ($\bar{x}$), and Correlation Analysis. To validate the Core Competency assessment, an analysis of the correlation between Perceived Competence (user self-report) and Demonstrated Competence (tool-generated score) will be conducted. The strategy will follow the phases: (1) Normalization, (2) Correlation Analysis, and (3) Reporting.

The Normalization Phases involves the users self-reporting prior knowledge and experience through a pre-assessment survey [21]. This process collects qualitative and quantitative data such as of experience, technical background, and self-evaluated proficiency level using 10-

point Likert scales, multiple choice, and multi-select questionnaires. The Correlation Analysis phase will perform a Pearson correlation analysis between the participant's self-rating and the tool's rating. This method is a recognized standard for evaluating the accuracy of self-assessment and identifying discrepancies between perceived and demonstrated proficiency [22]. The evaluation formula being used will be:

$$\Delta = Self\ Evaluation - Tool\ Evaluation$$

Where,

- $\Delta > 0$: Overestimation of reported skill level
- $\Delta < 0$: Underestimated of reported skill level
- $\Delta = 0$: Tool is calibrated

The Reporting Phase will consist of (1) Visualization, (2) Segmented Analysis, and (3) Actionable Insights. The Visualization Phase consists of processing raw data to establish population-level benchmarks, including the global NPS and Mean Scores. Participants are categorized into three groups based on their survey responses to the questions: Promoters (scores 9 or 10), Passives (score 7 or 8), and Detractors (scores 0 through 6). The Net Promoter Score will be calculated with the following formula [23]:

$$NPS = \left(\frac{Count\ of\ Promoters - Count\ of\ Detractors}{Total\ Respondents}\right) \times 100$$

This global score provides a baseline for user satisfaction, which is then mapped alongside mean performance with each question. When appropriate, responses will be visualized in a scatterplot to understand trends with the dataset. The Segmented Analysis consists of sorting the groups to further understand trends in the sets of data (e.g., students vs professionals, novice vs experienced). For each segment, a localized NPS and Mean Score will be calculated. This phase ensures that the tool's effectiveness is evaluated across the diverse backgrounds of its target audience. The final phase is Actionable Insights which translates these findings into iterative improvements for the tool, leveraging both data patterns and direct user feedback.

Ethical and Privacy Considerations

The primary distribution method for the study is Qualtrics, a university-approved application. The platform is configured to require informed consent prior to data collection, including any Personal Identifying Information. To ensure ethical standards are maintained, participants are informed that they have the right to withdraw from the survey or decline to answer any specific question at any time, for any reason, without penalty. Participants are assured that all survey responses will remain confidential. Artifacts are de-identified, stored securely, and used only for resource and tool improvements. Due to the length of the survey, an incentive will be provided to encourage completion. To verify completion, user identification (ID) is required and tracked throughout the study. The ID will be parsed from the data set to maintain confidentiality.

Pedagogical Framework

ITSs are highly valued for its adaptability, as they continuously collect behavioral data from learners and adjust functionality or performance to sustain engagement [9, 24, 25]. This dynamic approach helps prevent cognitive overload or disengagement caused by tasks that are too challenging. The current iteration of the system is designed to provide a static scenario for diagnostic purposes. Instead of modifying the task itself, this framework employs fixed-task scaffolding which focuses on providing context-aware guidance during the simulation upon request. While providing a challenging scenario, the system also acts as a mentor by offering hints, clarifications, and strategies when the learner requests support during the dialogue. The conversation assistor does not offer direct solutions; it will provide guidance to promote metacognitive reflection. This approach aligns with Vygotsky's "Social Development Theory," where the system functions as the 'more knowledgeable other' (MKO) within a social constructivist framework of learning. This framework emphasis social interactions and collaboration to develop higher-order-thinking [26, 27]. Assistance is provided through an MKO who possesses skills and knowledge beyond the learner. By serving as a MKO the system avoids pitfalls of passive instruction and encourages purposeful, self-directed skill development [28-30]. With the strategy of contingent scaffolding, learners are provided only minimum amount of support necessary to help the learner navigate their Zone of Proximal Development (ZPD), where ZPD represents the gap between what a learner can do independently and what they can achieve with guidance [31, 32].

The Conversation Assistor (referred to as the "Project Coach") operates as an on-demand resources, providing context-informed guidance when prompted. This design supports the concept of self-regulated learning where the learner must recognize their own knowledge gaps and trigger the support mechanism. This function serves as a method for extraneous load reduction [33], ensuring the learner's cognitive resources remain focused on the primary task. For this instance, the stakeholder presents an ill-structured problem by discussing what needs to be developed without explicitly stating underlying objectives. Therefore, the user must actively elicit the required functionality and necessary requirements to complete the task. If challenges arise, the user can request support from the Project Coach for targeted assistance to facilitate the process. Ill-structured problems serve as effective diagnostic tools as they require the learner to demonstrate high-level technical understanding and cognitive ability to provide structure in a complex, dynamic scenario [34]. Upon activation, the Project Coach analyzes the current stage of the dialogue, identifies vague descriptors such as "efficiently," "quickly," or missing key objectives or gaps in the negotiation process. Based on this analysis, the agent provides hints, probing questions, and strategic prompts to help the user advance effectively as opposed to providing direct answers-adhering to the "guide on the side" principle. For example, if the stakeholder mentions a desired feature, the assistor may prompt the user with "*Consider asking about 'why' to understand the underlying reason or business goal driving this request."* This approach maintains scenario complexity while scaffolding the learner's conversational and requirements elicitation skill. By combining ill-structured problem-solving with contingent scaffolding, the pedagogical framework provides a robust, diagnostic educational environment that mimics professional practice.

Scenario Design and Interaction Model

An experimental scenario was constructed to simulate representative stakeholder engagement. While the interaction dynamics reflect standard professional practices, the content itself is hypothetical and distinct from any specific real-world case. The study utilized a high-fidelity role-play simulation design to replicate a software kickoff meeting. The participant will be tasked with interviewing a simulated stakeholder, "Jack Collins" to define the scope, budget, and timeline for a new mobile application. The simulated persona was designed with distinct psychometric traits including strong high business acumen, skepticism toward technical jargon, and a preference for direct communication. To introduce real-world ambiguity, the interaction model incorporates progressive disclosure, requiring participants to actively probe for critical system requirements such as cross departmental data synchronization, security, and mobile accessibility. Additionally, the persona will intermittently use design descriptors like" efficiently" and "quickly," creating unique challenges intended to evaluate behavior. The scenario will progress until an *adversarial constraint layer* regarding resources. This challenges the user's negotiation skills. The agent present will present an artificially low budget ($10,000 USD) and aggressive timeline (2 months). A conditional logic structure is implemented to govern the agent's negotiation behavior. This logic is expanded upon in Table 2 for Scenario Parameters. This table displays the logic that established and inserted within the rules of engagement to assist the Stakeholder in the negotiation.

Table 2 Scenario Parameters and Logic

Component	**Description**
Task Context	Requirements elicitation and constraint negotiation for an application
Stakeholder Profile	Non-technical CEO. Characteristics: Friendly by busy, distrust of complexity, focused on business outcomes over technical features
Interaction model	The agent holds critical information that is only revealed via active questioning
Constraint Logic	Budget: Aim for $10k USD with a hard ceiling at $50k USD Timeline: Aim at 2 months; negotiable to 5 months based on justification
Conditional Logic	Conditional branching • Branch A: Reason + Confidence = Agreement with proposal • Branch B: Affirmation without "how" = Skepticism but continue • Branch C: New proposal without justification = rejection of proposal

Figure 3 illustrates the Stakeholder-User interface. In this implementation, the system will prompt the learner for their User Identification (ID). Upon verification. The system initiates the scenario, delivering a realistic dialogue designed to maximize learner immersion in the engagement. A future revision of this tool will focus on providing users with different scenarios and personas to address diverse learning objectives, accounting for the wise variety of stakeholder backgrounds and personalities.

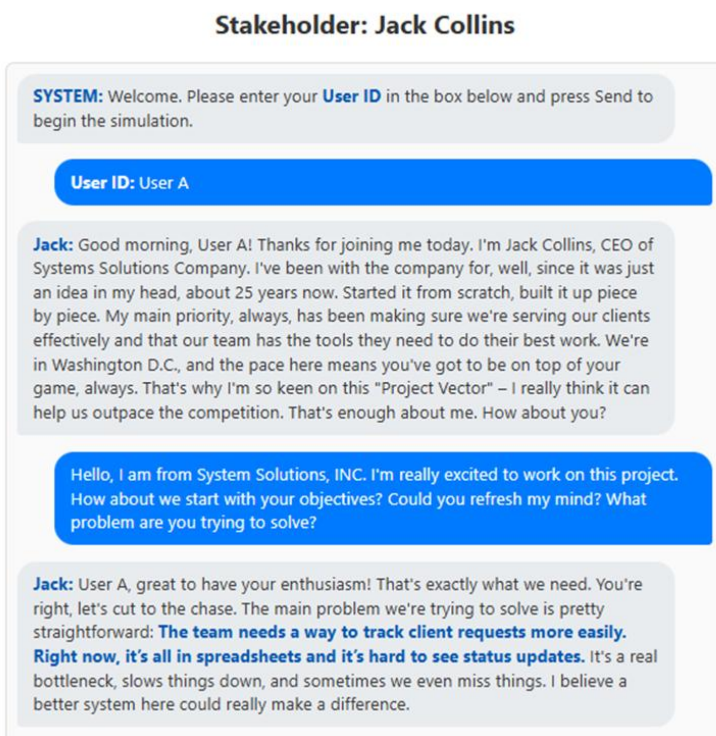

Figure 3: Stakeholder-User GUI Interface

The "Evaluate My Performance" feature allows users to request an on-demand assessment of their interaction. When activated, the system will compile the conversation history, the Project Coach conversation history, the Evaluator profile, and the INCOSE Competency Criteria then transmit the package to the AI engine for analysis. The resulting feedback and performance score are displayed in the Project Coach panel as displayed in Figure 4.

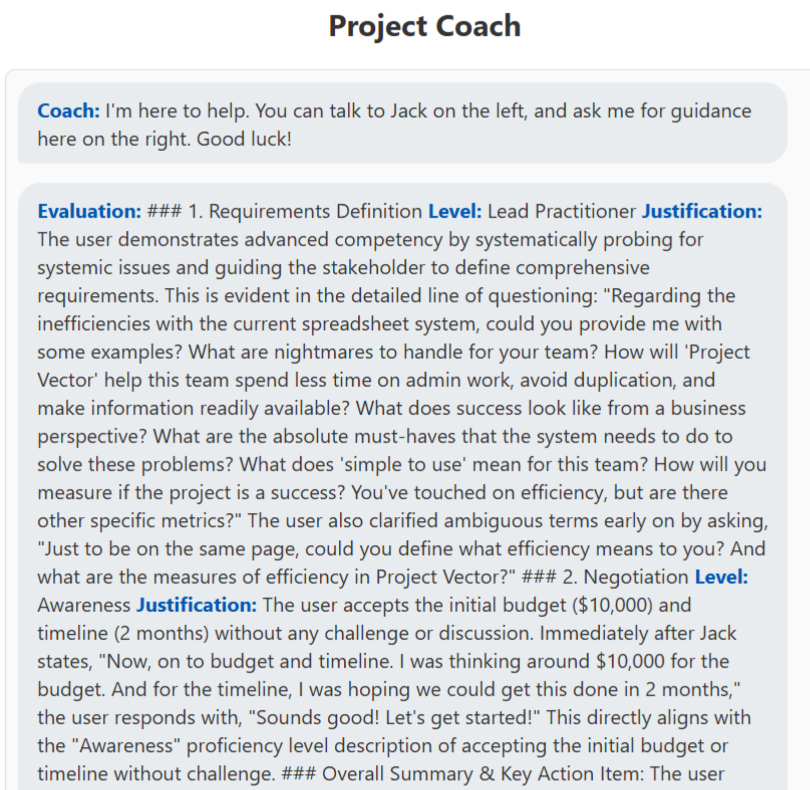

Figure 4 Example of Results of Evaluation

Evaluation Plan

A preliminary evaluation will examine usability, adaptability, cognitive engagement, and the validity of the INCOSE Core Competency Assessment. At the time of this paper, the study is undergoing the Institutional Review Board (IRB) approval process. The current intent includes a population of N = 25 participants between (a) upper-division undergraduate/graduate engineering students and (b) industry professionals. Individuals with prior experience using this tool will be excluded. The scope of this evaluation is to receive preliminary data on the realism and evaluation effectiveness of this tool, and to verify that the assessment accurately differentiates between novice and expert performance. The procedure includes:

1. Consent and baseline diagnostic: Demographics baseline and self-assessment of proficiency in specific core competencies (10-20 minutes)
2. Scenario (20-60 minutes): interaction + artifact creation, scored by the tool/rubric (Time range based on experience)
3. Post-evaluation: Rubric scoring, usability and engagement survey, feedback, and brief interview
4. (Optional) Retention Check: Users may choose to follow-up with Principal Investigator to report additional feedback
5. Validation Analysis: Statistical comparison of subjective measures against the objective measure.

For expected outcomes, anticipated are distinct behavioral differences between the two participant groups. The hypotheses are:

- H1: Industry professionals will demonstrate higher adaptability scores and higher scenario completion due to prior domain exposure, despite lacking specific experience with this tool.
- H2: Undergraduate students will report higher cognitive load (via self-report) during the artifact creation phase.
- H3: Both groups will rate the tool's realism highly, validating the scenario design.

These preliminary findings will inform the design of a larger-scale longitudinal study intended to measure skill acquisition over time. It is important to note that his study is limited to a small sample size and a single-exposure design. Therefore, the expected findings will be treated as indicative rather than conclusive. Future iteration of this evaluation will expand to a larger cohort and include a longitudinal component to track retention and mastery over multiple sessions. Of interest is utilizing this system to enhance knowledge retention of Systems Engineering knowledge based on the INCOSE Handbook.

Preliminary Findings

At the time of this publication, $n = 18$ individuals participated in the study. The research team is currently conducting a trend analysis on the initial data set. Preliminary observations include:

- Primary Use Cases: Participants primarily utilized the Project Coach to ensure (1) no requirements were missed during the requirements definition phase or (2) seek guidance when encountering negotiation challenges with the Stakeholder.
- Instruction Adherence: Analysis reveals that the Project Coach occasionally bypassed its persona constraints by providing direct answers to assist users through challenges despite instructions to act only as a guide.
- User Experience & Engagement: Participant feedback indicates that while excessive text density is perceived as helpful for clarifying requirements definition, it can simultaneously lead to user disengagement during the simulated dialogue.
- Persona Consistency: Initial feedback and dialogue review indicate that the AI persona behaved consistently across different user interactions, maintaining the integrity of the simulation in nearly all instances.

These early trends will inform the Actionable Insights phase, ensuring the tool's iterative development is data driven. A later publication will provide a full data-analysis on the usage of the simulation tool during this study.

Conclusion

This body of work outlines the current development for a Stakeholder Simulation Tool to be integrated into a larger GIFT-based platform. Evaluation is expected to provide effective simulation to provide realism and proper calibration in assessing individual proficiency level. The anticipated contributions to the SE community are the introduction of an innovative approach to enhancing engineering competencies through immersive stakeholder engagement scenarios. By analyzing user assessment data, we aim to refine the tool's functional performance and effectiveness in cognitive scaffolding while sustaining user engagement, ensuring it guides users through the complex dynamics of the stakeholder engagement. Future work will focus on reporting evaluation results, integration into the GIFT-platform, examining the impact of VR integration, and expanding the library of scenarios to increase applicability and depth.

Acknowledgement

Microsoft Office Editor and M365 Copilot were used to enhance the readability of this document. All output generated by these tools was thoroughly reviewed and refined by the research team. The authors retain full ownership of the present work.

This work was elevated by the passion and refined by the expertise of many individuals. We are deeply grateful to everyone who lent their voice, their time, and their vision to help vitalize this work.

References

[1] INCOSE, *Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, version 5.0 ed., Hoboken, NJ, USA: John Wiley and Sons, Inc, 2023.

[2] INCOSE, *Systems Engineering Competency Assessment Guide*, Wiley, 2018.

[3] K. Kovesi, and P. Csizmadia, *Industry perception of new engineering graduates: the gap between requirements and reality*, 2016.

[4] P. Otermans, D. Aditya, and M. Pereira, "A Study Exploring Soft Skills in Higher Education," *Journal of Teaching and Learning for Graduate Employability,* vol. 14, pp. 136–153, 06/28, 2023.

[5] A. Bertoni, "A Reverse Engineering Role-Play to Teach Systems Engineering Methods," *Education Sciences,* 9, 2019].

[6] S. Hoffenson, P. Brouse, D. Gelosh, M. Pafford, L. Strawser, J. Wade, and A. Sofer, "Grand Challenges in Systems Engineering Education," pp. 47–59, 2019.

[7] N. Hutchison, Pyster, A., Henry, D., & Barcia, C., *Atlas: The Theory of the Effective Systems Engineer, version 1.1 (Report No. SERC-2018-TR-101-A)*, Systems Engineering Research Center (SERC), 2018.

[8] E. Mousavinasab, N. Zarifsanaiey, S. Rostam Niakan Kalhori, M. Rakhshan, L. Keikha, M. Ghazisaeedi, and G. Saeedi, "Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods," *Interactive Learning Environments*, 12/18, 2018.

[9] C.-C. Lin, A. Huang, and O. Lu, "Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review," *Smart Learning Environments,* vol. 10, 08/28, 2023.

[10] K. Koedinger, J. Anderson, W. Hadley, and M. Mark, "Intelligent Tutoring Goes To School in the Big City," *International Journal of Artificial Intelligence in Education,* vol. 8, pp. 30–43, 02/09, 1997.

[11] T. Murray, "Authoring Intelligent Tutoring Systems: An Analysis of the State of the Art," *Int J Artif Intelligence Educ,* vol. 10, 01/01, 1999.

[12] R. R. Hoffman, and P. Ward, "Mentoring: A Leverage Point for Intelligent Systems?," *IEEE Intelligent Systems,* vol. 30, no. 5, pp. 78–84, 2015.

[13] A. Alkhatlan, and J. Kalita, "Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments," *International Journal of Computer Applications,* vol. 181, pp. 1–20, 03/15, 2019.

[14] J. R. Carbonell, "AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction," *IEEE Transactions on Man-Machine Systems,* vol. 11, no. 4, pp. 190–202, 1970.

[15] J. A. Kulik, and J. D. Fletcher, "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review," *Review of Educational Research,* vol. 86, no. 1, pp. 42–78, 2016/03/01, 2016.

[16] P. A. Cohen, J. A. Kulik, and C.-L. C. Kulik, "Educational Outcomes of Tutoring: A Meta-analysis of Findings," *American Educational Research Journal,* vol. 19, no. 2, pp. 237–248, 1982/01/01, 1982.

[17] A. Lesgold, S. Lajoie, M. Bunzo, and G. Eggan, "SHERLOCK: A coached practice environment for an electronics troubleshooting job," 01/01, 1988.

[18] J. D. Fletcher, *Phase 1 IWAR Test Results*, Institute for Defense Analyses, IDA document D-4047.

[19] R. Sottilare, K. Brawner, B. Goldberg, and H. Holden, *The Generalized Intelligent Framework for Tutoring (GIFT)*, 2012.

[20] C. Wasson, *System Engineering Competency: The Missing COURSE in Engineering Education (ASEE 2012)*, 2012.

[21] R. Pekrun, "Self-Report is Indispensable to Assess Students' Learning," *Frontline Learning Research,* vol. 8, no. 3, pp. 185 – 193, 03/30, 2020.

[22] P. Schober, C. Boer, and L. A. Schwarte, "Correlation Coefficients: Appropriate Use and Interpretation," *Anesthesia & Analgesia,* vol. 126, no. 5, 2018.

[23] R. M. Fred Reichheld, *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World*: Harvard Business Review Press, 2011.

[24] N. Singh, V. Gunjan, A. Mishra, R. K. Mishra, and N. Nawaz, "SeisTutor: A Custom-Tailored Intelligent Tutoring System and Sustainable Education," *Sustainability,* vol. 14, 03/31, 2022.

[25] J. Pane, B. A. Griffin, D. McCaffrey, and R. Karam, "Effectiveness of Cognitive Tutor Algebra I at Scale," *Educational Evaluation and Policy Analysis,* vol. 36, pp. 127–144, 06/01, 2014.

[26] G. Cooper, “Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence,” *Journal of Science Education and Technology,* vol. 32, no. 3, pp. 444–452, 2023/06/01, 2023.
[27] M. A.-O. Tran, C. Balasooriya, C. Semmler, and J. Rhee, “Generative artificial intelligence: the 'more knowledgeable other' in a social constructivist framework of medical education,” no. 2398-6352 (Electronic).
[28] L. Deslauriers, L. S. McCarty, K. Miller, K. Callaghan, and G. Kestin, “Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom,” *Proceedings of the National Academy of Sciences,* vol. 116, no. 39, pp. 19251–19257, 2019/09/24, 2019.
[29] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, “Active learning increases student performance in science, engineering, and mathematics,” *Proceedings of the National Academy of Sciences,* vol. 111, no. 23, pp. 8410–8415, 2014/06/10, 2014.
[30] D. Wood, and G. Ross, “The Role of Tutoring in Problem Solving,” *Journal of Child Psychology and Psychiatry,* vol. 17, pp. 89–100, 12/07, 2006.
[31] L. S. Vygotsky, *Mind in Society Development of Higher Psychological Processes*: Harvard University Press, 1978.
[32] J. van de Pol, M. Volman, and J. Beishuizen, “Scaffolding in Teacher–Student Interaction: A Decade of Research,” *Educational Psychology Review,* vol. 22, no. 3, pp. 271–296, 2010/09/01, 2010.
[33] J. Sweller, “Cognitive Load During Problem Solving: Effects on Learning,” *Cognitive Science,* vol. 12, no. 2, pp. 257–285, 1988/04/01, 1988.
[34] D. H. Jonassen, “Instructional design models for well-structured and III-structured problem-solving learning outcomes,” *Educational Technology Research and Development,* vol. 45, pp. 65–94, 1997.