Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 178–188
https://doi.org/10.54941/ahfe1006830

Standardizing Human Performance Evaluation in MBSE

Faysal Ahmmed and Sean Walker

Systems Engineering Program, University of South Alabama, Mobile, AL 36688, USA

ABSTRACT

This review examines research on the standardization of human capability, efficiency, and performance in model-based systems engineering (MBSE). The authors address fragmentation and inconsistency for integrating human factors in MBSE. The literature review aims to evaluate the current integration of human metrics in MBSE, explore modeling methodologies, and validation approaches. It also identifies the pathway toward standardization. On the other hand, due to a lack of real-time data, validation is also difficult for this process. Under these circumstances, standardizing human factors will open a new era in this field. It will lead to better, more accurate system designs and improve communication among human-digital components in a complex system.

Keywords: Model-based systems engineering (MBSE), Human factors, Human performance evaluation, Human-centered system (HCS), Human system integration (HSI), Standardizations

INTRODUCTION

Complex socio-technical systems require an engineering approach that includes both the performance of human operators and hardware, software components. Model-Based Systems Engineering (MBSE) has become an impactful method for managing complex socio-technical systems. This provides structured approaches for system design, analysis, validation, verification, and life cycle management (Zou et al., 2020). Using MBSE, a system engineer can create blueprints for every technical portion of a system. However, these models have a blind spot. They can't capture human elements such as an operator's performance factors such as response time, decision-making capability, physical limit, or stress status (Zou et al., 2020). As a result, integrating Human performance evaluation within the MBSE framework remains challenging and lacks standardization. This makes a big gap to establish a digital twin for an Industry 5.0 factory, where the success of the industry not only depends on its technology but also on its ability to optimize human performance (Adel, 2022). On the other hand, the proper validation depends on the standardization of human performance. For immersive technologies like Virtual humans and virtual patients, their success depends on replicating human attributes (Khan et al., 2023; 2024). Without a standardized benchmark for human behaviour, it is not possible to verify systems that replicate the behaviours of real humans.

HUMAN FACTORS

Human factor refers to the various human elements, such as cognitive, physical, psychological, thinking, emotional, and social characteristics of systems (Shih, 1999). These elements influence the system's performance and are influenced by the system and digital environment. In the aspect of engineering, Human factors address how people interact with the system, and what are impacts of various human elements, such as decision-making, workload, and ergonomics on the system. The primary aim of this concept is to optimize the system's design, increase efficiency in operation, and make it more user-friendly to humans. In a broader sense, these human factors play very important role to system's overall performance.

HUMAN-CENTRIC SYSTEM (HCS) AND HUMAN SYSTEMS INTEGRATION (HSI)

A human-centric system is defined as a system that is designed to work with people (Li & Xu, 2021). The main goal of this system is to create a positive and helpful relationship between human and machine rather than replacing humans. By understanding human language, emotion, and behaviour, this system makes humans' interactions with technology more effective and natural. The key characteristic of the system is to prioritize humans first (Li & Xu, 2021).

Human System Integration (HSI) is a way of designing a system that focuses on the people who use the system. This concept aims to integrate human, organizational, and technical aspects throughout a system's lifecycle (Disdier et al., 2024). In traditional systems engineering, human and organizational stakeholders are treated as entities that interface with the system. Whereas HIS adopts a sociotechnical systems perspective. In HSI, humans are considered a system within a larger system of systems. In this aspect, human factors are not considered as external elements but are embedded in the architecture of the systems (Disdier et al., 2024).

PROBLEM STATEMENT

For designing an efficient and effective socio-technical system, it is essential to represent human factors, those that are involved throughout the process, accurately. In a socio-technical system, when we include evaluation metrics for human operators such as decision-making speed, task completion, response time, and error rate, that system can function effectively to serve the purpose. This leads to clearer roles, reduced operational risks, and gives a holistic overview of the system. However, engineers are facing major challenges that include inconsistency and fragmentation. Research to integrate human performance is often scattered (Sarathi et al., 2022). Different teams develop their methods separately, without a common framework, and as a result, these outputs do not fit together. This fragmentation makes it difficult to connect human capability to the system's requirements holistically. The problem becomes worse in the absence of standard metrics and modeling techniques for human factors in a socio-technical system. Due to the absence of standard metrics, each project often creates a custom solution, and these solutions cannot be reused in the future. On the other hand, the whole process raises a question about the validation process. It is very difficult to get real-time data from an actual operational environment to test the model. Without proper validation, the idea of a human performance model remains theoretical rather than becoming a practical, reliable tool. Therefore, standardizing the evaluation of human performance within MBSE is an engineering necessity. Research bridges the gap between Human System Integration (HSI) in various cutting-edge fields, such as digital twins for Industry 5.0 and Immersive technology.

Objectives

By conducting a comprehensive review and analysing the existing state-of-the-art, this paper aims to contribute to this emerging field. The primary objectives are: a) To explore current knowledge on integrating human factors and performance metrics within MBSE models, b) To explore tools and methods used to define the human performance metrics, c) To explore approaches for validating human performance models embedded in MBSE frameworks, and d) To study challenges and propose pathways toward standardizing human performance metrics in MBSE environments.

METHODOLOGY

Review protocol: PRISMA guidelines (Page et al., 2021). The PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) is an internationally recognized framework for reviewing literature. Providing a checklist and flow diagram ensures the transparency and the quality of reporting for systematic review and meta-analysis.

Search strategy: To identify relevant articles, a comprehensive search strategy is developed. This search is designed to capture research on MBSE, human factors, human performance, human system integration (HSI), and standardization.

Sources: WILEY Online Library, ScienceDirect

Search string: (MBSE) AND (Human factor OR Human performance OR Human capability OR Human System Integration OR Standardization)

Date range: Studies published between 2020 to 2025

Screening process: Screening through “Title” and “Abstract”

Inclusion Criteria:

a) Article published between 2020–2025.
b) Study focusing on MBSE.
c) Studies addressing human factors, human performance, human capability, human system integration, and standardizing human factor metrics.
d) Peer-reviewed articles, conference papers, and journals.
e) Studies available in the English language.

Exclusion criteria:

a) Abstract only publications, editorials, commentaries, and opinions.
b) Duplicate publications.

Data extraction and analysis: After the screening process, data is extracted and analysed with a focus on key contents of the article, the method used, the tools used, parameters used to define human factors, metrics to measure performance, and validation status.

Analysis and Synthesis of Findings

This literature review explores the current state of standardizing human performance evaluation within MBSE. Reviewing the selected studies expresses the ongoing transition of this field. The findings from the analysis represent the main research objectives as below:

Table 1: Summary of recent literature on human factors in model-based systems engineering.

Author, Year	Key Content	Method Used	Tools Used	Parameter Standardization	Validation Status
(Sarathi et al., 2022)	Human performance, Human traits, states and skills (TSS)	MBSE, Function Behavior Structure (FBS)	MBSE tool (nothing specific, could be CAMEO, SIMIO)	Cognitive fatigue, stress level, vision acuity, physical, cognitive, social, and emotional domains	Conceptually Validated.Not yet validated with a complete, real-world mission dataset or real-time data.
(Boy et al., 2024)	Comprehensive methodology for integrating and validating human performance	PRODEC	BPMN for procedural modeling, iBlocks formalism for task/activity analysis. HITLS facilities & Digital Twins, Wizard-of-Oz prototyping.	SADMAT: Situational-Awareness, Decision-Making, Action-Taking, plus workload, trust, collaboration; Human Readiness Levels (HRL), Organizational Readiness Levels (ORL) and Technology Readiness Levels (TRL)	Validated via Case Study

(Continued)

Table 1: Continued

Author, Year	Key Content	Method Used	Tools Used	Parameter Standardization	Validation Status
(Birch et al., 2023)	Proposes integrating a Human Factors Hazard Model (HFHM) into SysML for Model-Based Systems Engineering (MBSE)	Human Reliability Analysis (HRA) Fault Tree Analysis (FTA) Event Tree Analysis (ETA)	SysML, Risk Analysis and Assessment Modeling Language (RAAML) profile, Cameo Systems Modeler software, Microsoft Excel	Risk Analysis and Assessment Modeling Language (RAAML) standard	Verified and Validated Verification: Compared against the THERP method; average 95.2% agreement Validation: Conducted via a sensitivity trade study showing logical results and a Subject Matter Expert (SME) review (average Likert score of 3.5/4.0)
(Corl & Gallegos, 2024)	Introduces Relational and Technological Capstone (RTC) framework for HSI in MBSE, Integrates human factors into SysML models	Two-phase RTC method, Performance aggregation using weighted harmonic mean, Performance Shaping Factors (PSFs),	Cameo, SysML, Python, Simulation Toolkit plugin	10 standardized PSFs (e.g., Task Training, Mental Acuity, Stress)	Validated via parametric simulation in Cameo, Case study applied to UUV operator in SDV context. No real-time validation or external field testing
(Peleg, Gavish & Zonnenshain, 2022)	Proposal for a holistic MBSE methodology to support HSI throughout the system lifecycle. Integration of HSI into a Digital Twin knowledge base.	Metaphor Vision MBSE (MV-MBSE).	Sparx Enterprise Architect (EA), Metaphor Builder (MB) Domain Specific Languages (DSL)	Process-time, resource use, correctness via simulation	Validate through 3D visual simulation. No real-time data validation or field testing

(Continued)

Table 1: Continued

Author, Year	Key Content	Method Used	Tools Used	Parameter Standardization	Validation Status
(Gräßler, Wiechel & Roesmann, 2021)	SysML meta-model extension for human factors. Integration of human agents with individual characteristics.	5-step MBSE approach	Systems Modeling Language (SysML). Unified Modeling Language (UML)	Mental/cognitive factors: Skill Level (Novice, Advanced Beginner, Competent, Proficient, Expert), IT Knowledge, language skills, experience. Physical factors: arm length, vision capabilities, handedness	Validated through case study

To integrate human performance, some researchers have focused on creating structured architectures. Sarath et al., developed a MBSE-Systems Architecture (MBSE-SA) to evaluate the use in case of various missions for the Department of Defence (DoD) (Sarathi et al., 2022). This model captures various human capabilities as a standardized taxonomy of Traits, States, and Skills (TSS). By structuring these human factors, Sarathi et al. created a common language in which these human factors are integrated through a formal conceptual architecture. This framework emphasizes approaching the problem from a broader view rather than looking only at a single system or project. Systems Architecture (SA) is designed to build a framework that makes connections clear and traceable. From this point of view, integrating human factors into a system is not about joining the technical model and the human factors. Instead, it is about addressing relationships across multiple layers, such as the objectives, the environments, the capabilities of humans, etc.

A common pattern was found in how researchers have incorporated the human elements into SysML-based system models. For example, Peleg et al defined human actor according to the process they are involved in the system (Peleg et al., 2022). Furthermore, Gräßler et al represented people through dedicated Human Agent stereotypes, which gave them a formalized place in the system as system components (Gräßler et al., 2021). According to Birch et al, the human is a reliability component that counts as a formal element in Sys-ML model (Birch et al., 2023). All of these approaches define that Human Factors are not an external or informal consideration but a formal modeling element. This formal element can be structurally integrated into Sys-ML diagrams. As a result, it becomes possible to connect roles to system tasks, define attributes to human actors, and analyse the impact of human factors more systematically. Beyond this foundation, Corl and Gallegos take the concept further by developing a human performance measure into the model. More specifically they introduced the parameter P_RTC as a constraint property to measure human performance of a System

(Corl & Gallegos, 2024). This means human performance is no longer just described in abstract terms but actively computed in a system.

A holistic view is created with the PRODEC methodology for Human-Machine Integration (HMI) (Boy et al., 2024). This offers a human-centric approach to thinking about how humans and machines work together. In this methodology, both humans and machines function as cognitive and physical agents. These agents interact continuously and shape one another over time. In this framework, operators bring practical knowledge from their experience to operate the machines. On the other hand, the system designer contributes to formalizing or formulating rules, models, and structured information about the systems. Here, Digital Twin plays a crucial role by acting as a medium where the two forms of knowledge are brought together and continuously exchanged. From this viewpoint, Human performance is not just the sum of individual abilities but an emergent technology that grows out of continuous collaboration of Human and machine.

DISCUSSION

Human factors are shifting from simply representing humans in a system to a real operational state in SysML models. Therefore, this research found a noticeable divide in the present research approach. On one side, some researchers emphasize creating structured taxonomy and standards (Gräßler et al., 2021; Sarathi et al., 2022). This defines the aspect of human work and behavior that should be represented in the system. We found a common goal here, consistency and comparability, which ensures the existence of the human factor in a unified way across different systems. On the other hand, researchers like Peleg, Boy and Corl are more interested in the dynamic side of the problem. They focus on the process to evaluate human performance through simulation. This method enables experimentation with human-system interaction in real and hypothetical scenarios for measuring outcomes and identifying risks and improvements (Peleg et al., 2022; Boy et al., 2024; Corl & Gallegos, 2024). Both of these approaches are complementary to each other. The most promising path forward would be to bring them together under the same umbrella by linking standardized human ontologies with a simulation environment that enables consistent representation and practical evaluation of human performance in complex systems.

The current research shows that there is no single, universal tool for integrating human performance in MBSE. For the methodological landscape, the best way would be to develop an MBSE toolbox rather than a single standardization approach. Architecture infrastructure can be developed through core MBSE platforms such as Cameo System Modeler and Sparx EA (Gräßler et al., 2021; Peleg et al., 2022; Birch et al., 2023; Corl & Gallegos, 2024). On the other hand, a range of modeling methods, such as process modeling, parameter modeling, and taxonomy development, define human tasks and attributes in complex Socio-Technical systems. For example, metrics introduced by Peleg and Boy appear as an implicit output solution, while Sarathis et al worked on analytical traceability. In both cases, metrics are not fully standardized. The notable work is done by Corl and Gallegos on performance metrics. They developed a formula for embedding metrics directly into SysML with calculable properties (Corl & Gallegos, 2024). This highlights the current gap and opens a direction for moving forward with more rigorous standardized human performance evaluation.

Absence of a universally accepted taxonomy for quantifying human performance parameters to be embedded with MBSE seamlessly is identified as a critical limitation of the literature review. Though existing top-level frameworks such as TSS and SADMAT provide valuable structure for gathering knowledge, these frameworks are more specifically usable than a reusable common framework (Sarathi et al., 2022; Boy et al., 2024). In the absence of a standardized, reusable library, each project requires inventing a set of metrics or depends on implicit, simulated data. This lack of standardization makes the process difficult to establish a benchmark by comparing results across various projects from various domains.

In the aspect of validation of Human performance integrated with MBSE, we found two different approaches. The first one is empirical and dynamic, which focuses on the realistic interactions of humans in work. To establish this researcher chose interactive simulations. Peleg applies 3D visual simulation while Boy emphasizes the role of Digital Twin-based Human-in-the-Loop Simulation (HITLS) (Peleg et al., 2022; Boy et al., 2024). By engaging individuals with real-world operational knowledge, this approach becomes a Subject Matter Expert (SMEs). Interacting with simulations, SMEs can verify logical soundness, usability, and practical feasibility. This approach prevents models from becoming elegant but functionally irrelevant. In addition, dynamic validation offers analyses and comparative benchmarks. For example, Birch confirmed mathematical accuracy through comparing model outputs with the legacy Technique for Human Error-Rate Prediction (THERP) methodology (Birch et al., 2023). The second method is analytical, static validation. This method focuses on the coherence and completeness of the model as a decision-making tool. Sarathi evaluates her models by their ability to find structural gaps. For example, identifying critical human traits that are not included in system elements. Similarly, Gräßler et al. use structural compatibility checking, where a model's value is defined by its ability to match individuals' capabilities with requirements for different roles formally.

From the analysis, the authors found a solid foundation for model-centric verification for both dynamic and analytical validation. This is the next step of checking whether the model matches real-world behavior. Within this step, it verifies that the model is consistent, complete, and logically sound. The challenges and pathways toward standardization found in the literature are given below in Table 2.

Table 2: Challenges and proposed pathway from literature.

Challenges	Proposed Pathways From Literature
Lack of standardized metrics and methods	Develop HIS-Specific Domain-Specific Language (Peleg et al., 2022). Follow emerging standards like RAAML for safety (Birch et al., 2023).
Subjective and expert-centric practices	Structured taxonomies to convert subject assignment to objective data (Sarathi et al., 2022; Corl & Gallegos 2024).
Separate research work and lack of traceability	Develop a reusable and standardized model (Gräßler et al., 2021; Peleg et al., 2022; Sarathi et al., 2022).
Dynamic validation gap	Use Digital-twin as the standard platform for an activity-based platform (Boy et al., 2024). Fully integrated and performance calculation within MBSE (Birch et al., 2023; Corl & Gallegos, 2024).

Identified Research Gaps and Future Work

Based on the synthesis, several research gaps to evaluate human performance in MBSE are identified. Moving forward, these gaps open the door to opportunities for future investigation.

The most significant gap is the absence of standardization. There is no complete computable ontology for human performance metrics. Future researchers need to focus on defining a core set of quantifiable parameters (with units, data types) and metrics (how to measure the parameters). So that the output could be manifested in MBSE for further applications in different fields, such as a Digital twin for manufacturing plants or a Virtual Human in immersive technology. This research should be able to bridge the gap between top-level system architecture to low-level measurable data, such as response time, decision-making time, etc.

In the aspect of a complex socio-technical system, the models found from this review are not enough to manage complexity. They didn't define how human performance interacts with other system components. What would be the impact on overall system performance? Under these circumstances, going forward, researchers should focus on developing a complete framework that defines the interaction between all other components of the system with human factors.

Researchers should develop a domain-based library (Ex, Manufacturing, medical, education, defense, etc) to evaluate human performance. So that it can illustrate the system more accurately.

Though papers illustrate some works on human performance through structuring or simulation, a completely integrated framework is still missing. In the future, researchers should develop such a model that integrates the standardized human performance ontology with simulation. This is how the research would be more functional by defining what to measure and how it is measured through the simulation with real-time data.

CONCLUSION

This review shows that MBSE offers a significant way to standardize human performance evaluation. This is clear that shifting from traditional document-based HSI (human factor as external element) to model-based HSI has already started. Researchers have made some strong progress in developing platforms (architecture, modeling language, profiles, and methods) and processes (simulation, traceability analysis) for human-centric design. However, Parameters for the Standardization of the performance evaluation and its metrics remain a major challenge. Though there are some approaches, they exist separately which is an obstacle to connecting in a unified way. There needs to be more research to be done for true standardization. True standardization would enable Systems Engineers to use a standardized library where they can simply drag and drop validation human performance metrics For example, The Human Performance as a block in SysML when dragged from the library should connect automatically to the simulation engine for dynamic testing and for verification it should connect to the requirements block. This allows full traceability from requirements to simulation results in an MBSE environment. Achieving this vision will enable an integrated human performance evaluation in a complex socio-technical system.

REFERENCES

Adel, A., 2022, *Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas*, *Journal of Cloud Computing*, 11(1).

Birch, D. S., Narsinghani, J., Herber, D. & Bradley, T. H., 2023, 'Human factors hazard modeling in the systems modeling language', *Systems Engineering*, 26(3), 328–343.

Boy, G. A., Masson, D., Durnerin, É. & Morel, C., 2024, 'PRODEC for human systems integration of increasingly autonomous systems', *Systems Engineering*, 27(5), 805–826.

Corl, K. & Gallegos, E., 2024, 'A Method for Human Systems Integration Requirements within Model Based Systems Engineering', *INCOSE International Symposium*, 34(1), 1787–1806.

Disdier, A., Masson, D., Jankovic, M. & Boy, G. A., 2024, 'Defining and characterizing the operational context for human systems integration', *Systems Engineering*, 27(6), 1103–1115.

Gräßler, I., Wiechel, D. & Roesmann, D., 2021, *Integrating human factors in the model based development of cyber-physical production systems*, *Procedia CIRP*, vol. 100, 518–523, Elsevier B. V.

Khan, M. N. R., Ahmmed, F., Zabir, M. K. Al, Lohar, B., Walker, S. & Lippert, K. J., 2024, *Concept Exploration of Virtual Humans in Education*, *Proceedings of International Conference on Smart Systems and Technologies, SST 2024*, 95–100, Institute of Electrical and Electronics Engineers Inc.

Li, T. & Xu, G., 2021, 'Human-Centric Intelligent Systems: A Welcome Note from Editor (s)-in-Chief', *Human-Centric Intelligent Systems*, 1(1–2), 1.

Martins Shih, H., 1999, 'Human Factors in Systems Engineering', *Human Factors and Ergonomics in Manufacturing*, 9(1), 131–132.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P. & Moher, D., 2021, 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews', *BMJ*, 372.

Peleg, E., Gavish, N. & Zonnenshain, A., 2022, 'Model Based HSI & Operational Process Simulation', *INCOSE International Symposium*, 32(S1), 92–108.

Rahman Khan, M. N., Ahmmed, F., Zabir, M. K. Al & Lippert, K. J., 2023, *Virtual Patient in Medical Education*, *7th International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc.

Sarathi, T., Cyr, J., Morris, H., DeLaura, R., Balcius, J., Collins, P. & Shatz, M., 2022, 'Developing a Human Performance Model Based Systems Engineering System Architecture (MBSE-SA) for Defense Applications', *INCOSE International Symposium*, 32(1), 608–622.

Zou, M., Vogel-Heuser, B., Sollfrank, M. & Fischer, J., 2020, *A cross-disciplinary model-based systems engineering workflow of automated production systems leveraging socio-technical aspects*, *IEEE International Conference on Industrial Engineering and Engineering Management*, vols. 2020-December, 133–140, IEEE Computer Society.