

Multi-dimensional Research on Human-computer Interaction System Evaluation

Xiu Miao¹, Huihui Zhao^{1*}

¹ School of Architecture and Art Design, Inner Mongolia University of Science and Technology, Baotou, China

*Correspondence: Huihui Zhao

Received: 6 March 2026; Accepted: 27 March 2026; Published: 23 April 2026

Abstract:

Human-machine system evaluation is a key link in optimizing the interaction experience between humans and machines, improving system performance and user satisfaction. This article reviews traditional and modern assessment methods, analyzes key current challenges, and discusses future research directions. Future human-machine system evaluation will be more intelligent, precise and personalized, thereby improving system performance and user experience. Based on this research analysis, human-machine system evaluation was obtained as human-machine interaction achieves the goals of harmony and user satisfaction.

Keywords:

Human-Machine System Evaluation, Multimodal Data Fusion, Virtual Reality

1. Introduction

With the rapid advancement of technology, Human-Machine Systems (HMS) are being increasingly applied across various fields. These systems refer to collaborative frameworks where humans and machines (including computers, robots, and smart devices) interact to accomplish tasks. The design and optimization of such systems require not only technical innovation but also consideration of human users' needs, behaviors, and psychological characteristics. Technological progress in artificial intelligence, the Internet of Things, and big data has significantly enhanced the complexity and intelligence of HMS. Typical examples include autonomous vehicles, smart home systems, and medical robots. In terms of application expansion, HMS have evolved from traditional industrial control to encompass healthcare, transportation, military, and consumer electronics sectors. Each domain presents unique requirements and challenges. As user expectations for products and services continue to rise, system usability, safety, and efficiency have become critical factors, while User Experience (UX) has gained increasing prominence in HMS design.

The importance of human-machine system evaluation cannot be overstated. By enhancing system performance, assessments reveal bottlenecks and deficiencies for optimization. Evaluating autonomous driving systems' response times and decision accuracy improves safety and reliability. Assessments help designers understand user needs and behaviors, enhancing user experience through systems that better align with user habits and expectations. User testing identifies operational challenges in smart home systems, enabling interface improvements. In high-risk fields like healthcare,

aviation, and military, human-machine system safety and reliability are critical. Evaluations help identify potential risks and failures, allowing for preventive measures and contingency plans. Promoting technological innovation, assessments drive the development and application of new technologies. Evaluating user experiences in VR and AR systems facilitates their adoption in education and entertainment. Assessment results provide scientific basis for policymakers and business managers, supporting decisions on technology investment, product development, and market promotion. Social and economic impacts are also significant: optimized human-machine systems improve quality of life and social welfare. For instance, smart healthcare systems enhance diagnostic accuracy and reduce medical errors. Economically, efficient human-machine systems boost productivity and lower operational costs. Smart manufacturing systems minimize waste and optimize resource utilization. With the continuous advancement of technology and the expansion of application fields, research on human-machine system evaluation will face more challenges and opportunities. Future studies may place greater emphasis on interdisciplinary collaboration, integrating knowledge from psychology, sociology, engineering, and other disciplines to develop more comprehensive and systematic evaluation methods.

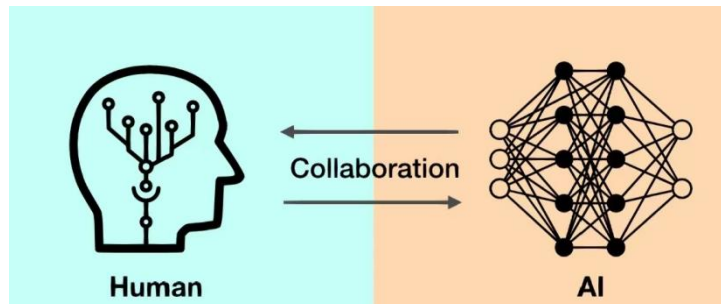


Figure 1. Collaborative Intelligence: Humans and AI Working in Tandem.

2. Man-machine Symbiosis

The fundamental concept of HMS is a system where humans and machines (including computers, robots, and smart devices) collaborate through interaction to accomplish tasks. The design and optimization of such systems require not only technological innovation but also consideration of human users' needs, behaviors, and psychological characteristics. [1] It primarily consists of human operators and machine equipment, achieving information transmission and processing through Human-Machine Interaction (HMI) to jointly accomplish specific tasks or objectives. [2] The core of human-machine systems lies in the collaborative work between humans and machines, emphasizing their complementary and interactive nature.

3. The Importance of Human-Machine System Evaluation

3.1. Enhance System Performance

Through evaluation, we can identify bottlenecks and deficiencies in the system for optimization and improvement. For instance, in autonomous driving systems, assessing response time and decision accuracy is critical. If the system's response time is too long, it may fail to make correct decisions promptly when encountering emergencies, potentially leading to traffic accidents. [3] By collecting and analyzing extensive real-world driving scenario data, as well as simulating various complex road

conditions for testing, we can accurately evaluate the performance of autonomous driving systems. [4] This process helps identify existing issues, enabling optimization of algorithms and sensor configurations to enhance safety and reliability.

3.2. Enhance User Experience

Evaluation helps designers understand user needs and behaviors, enabling them to create systems that better align with user habits and expectations. [5] Take smart home systems as an example: user testing can reveal operational challenges. For instance, some smart home devices have overly complex control interfaces, making it difficult for users to quickly locate required function buttons. [6] Additionally, insufficiently intelligent inter-device coordination settings may fail to meet daily usage requirements. Based on these insights, designers can refine interface designs, simplify operation workflows, optimize device integration logic, enhance user experience, and improve both satisfaction and efficiency.

3.3. Ensure Security and Reliability

In high-risk sectors such as healthcare, aviation, and military operations, the safety and reliability of human-machine systems are paramount. By evaluating potential risks and failure points within these systems, appropriate preventive and response measures can be implemented. In the medical field, rigorous assessment of medical devices is essential to ensure their stable and accurate operation under various complex clinical environments. Equipment failures may pose severe threats to patient safety. Through comprehensive evaluations of medical devices' performance, stability, and safety, potential risks can be identified in advance. This enables the development of maintenance plans and emergency response protocols, thereby effectively ensuring the safety and reliability of medical procedures.

3.4. Promote Technological Innovation

Evaluation studies can drive the development and application of new technologies. By assessing user experiences with virtual reality (VR) and augmented reality (AR) systems, researchers identify limitations in existing technologies regarding immersion and interactivity, which motivates them to develop new techniques and algorithms to address these issues. For instance, to enhance VR immersion, researchers continuously explore advanced display technologies and more precise position tracking systems. To improve AR interactivity, they develop new technologies such as gesture recognition and voice interaction. The application of these innovations further promotes the widespread adoption of VR and AR technologies in fields like education and entertainment.

3.5. Support Decision-making

The evaluation results provide scientific evidence for policymakers and business managers to support their decisions in technology investment, product development, and market promotion. Business managers can use these findings to assess whether a new technology holds market potential and justifies substantial R&D and production investments. Policymakers can also leverage the data to formulate relevant industrial policies, guiding the healthy development of industries. For instance, in emerging human-computer interaction technologies, policymakers may provide financial support and policy incentives to promising technologies based on evaluation results, thereby accelerating the growth of related industries.

3.6. Social and Economic Impacts

Optimized human-machine systems enhance social benefits, improve quality of life, and boost public welfare. By leveraging artificial intelligence and big data technologies, intelligent healthcare systems enable more accurate diagnosis and prediction of patient conditions, improving diagnostic and therapeutic precision while reducing medical errors to deliver better medical services. Efficient human-machine systems also drive economic growth by increasing productivity and lowering operational costs. Through automated production and smart management, intelligent manufacturing systems minimize waste in production processes, optimize resource utilization, reduce production costs, and strengthen corporate market competitiveness.

With the continuous advancement of technology and the expansion of application fields, research on human-machine system evaluation will face more challenges and opportunities. Future studies may place greater emphasis on interdisciplinary collaboration, integrating knowledge from psychology, sociology, engineering, and other disciplines to develop more comprehensive and systematic evaluation methods.

4. The Purpose and Core Problem of Man-machine System Evaluation

The primary objective of human-machine system evaluation is to ensure that systems meet expected performance, safety, and user experience requirements throughout their design, development, and deployment processes. Key objectives include: verifying system performance to assess whether it can efficiently and accurately accomplish predefined tasks while ensuring technical performance aligns with design specifications; optimizing user experience by identifying issues encountered during system use, refining interface design and interaction processes to enhance user satisfaction and operational efficiency; ensuring safety and reliability through risk identification and preventive measures to guarantee secure and dependable system operation under various conditions; supporting decision-making by providing scientific evidence for system designers, developers, and decision-makers to inform technical improvements, product optimization, and market promotion strategies; and fostering technological innovation by identifying shortcomings in existing systems, driving the development and application of new technologies, and promoting continuous innovation in the field of human-machine systems [7].

The core issues in human-machine system evaluation include: the efficiency, accuracy, reliability, and scalability of system performance; the usability, satisfaction, accessibility, and emotional experience of user experience; the failure rate, fault tolerance, data security, and emergency response of safety and reliability; the interface design, interaction process, and feedback mechanism of interaction design; and the physical, social, and technical environments of environmental adaptability.

5. Research Status

5.1. Traditional Evaluation Methods

5.1.1. Performance Appraisal

Task performance evaluation measures system performance based on the time, accuracy, and efficiency of task completion. Specifically, recording the time required

for users to complete tasks provides a clear understanding of the system's operational convenience.

5.1.2. Subjective User Evaluation

User subjective evaluation primarily collects users' perceptions of the system through questionnaires and interviews. Standardized questionnaires, such as the System Usability Scale (SUS) and the Questionnaire for User Interface Satisfaction (QUIS), can provide a relatively objective assessment of user satisfaction.

5.1.3. Assessment of Physiological Indicators

Physiological metrics evaluation assesses system usability by monitoring users' physiological responses. Eye-tracking technology records gaze movements, analyzes attention distribution patterns, and identifies key focus areas during interface interactions. Heart rate monitoring assesses the user's stress or fatigue level through changes in heart rate. Electroencephalography (EEG) analyzes users' cognitive load and emotional states through brainwave analysis. For instance, Guo Zhendong et al. described the design and implementation of an intelligent sleep aid device based on EEG signals. By real-time monitoring and analysis of users' brainwave signals, and using specific algorithms to adjust the output audio, the device can improve users' sleep quality [8].

5.1.4. Behavior Observation

Behavior observation is to evaluate the system design by observing the user behavior. For example, Li Jing et al. employed behavioral observation methods to investigate the transition intention of academic users from traditional to AI-powered literature reading tools, along with the behavioral influencing factors [9].

5.2. Modern Assessment Method

5.2.1. Eye-tracking Technology

Eye tracking technology is to record the user's gaze movement to analyze the distribution of attention and visual behavior. In web design, eye-tracking technology helps identify users' initial focus areas and content with prolonged attention spans during browsing, enabling optimized layout and information presentation. For advertising effectiveness evaluation, it analyzes user gaze patterns to assess ad appeal and performance. In user experience research, eye-tracking provides actionable insights for product design improvements.

5.2.2. Brain-computer Interface BCI

Brain-computer interface (BCI) assesses users' cognitive and emotional states through electroencephalogram (EEG) signals or other neural signals. In cognitive load assessment, EEG signal variations during complex task execution enable quantification of cognitive workload. For emotional response evaluation, specific EEG characteristics are analyzed to identify users' emotional states. In brain-computer interfaces (BCIs), this technology facilitates direct neural-machine communication, charting new course for future human-machine interaction development.

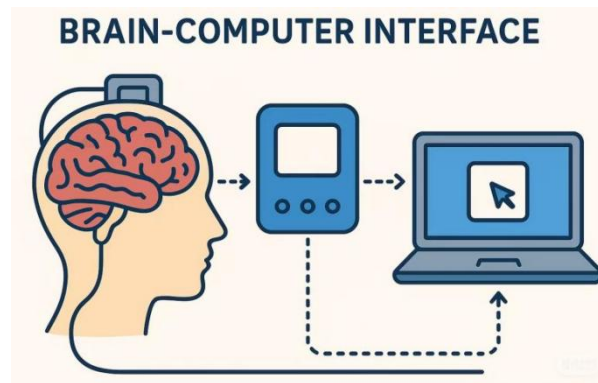


Figure 2. BCI Signal Acquisition and Feedback Process

5.2.3. Affective Computing

Emotional computing evaluates emotional states by analyzing users' facial expressions, voice, and physiological signals. Its primary application scenarios include emotion recognition, user experience research, and intelligent customer service. Among these, the market reputation evaluation based on e-commerce review text analysis by Wang Siwen et al. can provide crucial market feedback information to promote the development of Xinjiang fresh apricot e-commerce industry. This involves in-depth analysis and mining of review texts, including determining the emotional tendency of review corpora, extracting review topics based on LDA (Latent Dirichlet Allocation) topic models, and conducting quantitative consumer reputation evaluation through dependency syntactic analysis and emotional computing [10].

5.2.4. Virtual Reality (VR) and Augmented Reality (AR)

Virtual Reality (VR) and Augmented Reality (AR) evaluate human-computer interaction in virtual or augmented environments. In virtual training, VR technology creates realistic training scenarios to assess trainees' operational performance and learning outcomes. In game design, VR/AR experiences optimize interaction methods and narrative design. For medical simulation, these technologies replicate surgical scenarios to evaluate physicians' operational proficiency and accuracy in simulated environments.

5.2.5. Multimodal Assessment

Multimodal evaluation combines multiple data sources (e.g., eye movements, EEG, physiological signals) for comprehensive assessment. When evaluating user experience holistically, this approach captures user information from multiple dimensions, enabling a more thorough understanding of their needs and feelings. In complex system evaluations, integrating diverse data types enhances assessment accuracy and reliability. For instance, Zhang Ao et al. developed a comprehensive reliability evaluation method for complex systems that incorporates operational profile information, aiming to fully utilize reliability data accumulated across system layers and improve the effectiveness of reliability assessments. These modern evaluation methods, combining advanced technologies and interdisciplinary theories, provide more comprehensive and precise assessments of human-computer interaction effects, helping designers create systems that better meet user needs. [11].

6. Challenges and Problems in Man-machine System Evaluation

The challenges and issues in human-machine system evaluation are multidisciplinary and complex. Given the significant differences in user backgrounds, skills, and cultural diversity, a single evaluation method cannot be universally applicable. This necessitates the design of universal evaluation methods to ensure results are suitable for diverse user groups. The integration of multimodal data involves multiple data types, making integration and analysis complex. Effective multimodal data integration is crucial to provide comprehensive evaluation results. For instance, Tian Chenzhi et al. introduced eye-tracking technology, utilizing pupil diameter, eye movement frequency, and fixation point counts to construct a quantitative evaluation model for command and control system human-machine interaction. This model assessed search efficiency, operational efficiency, cognitive load, design complexity, and attractiveness, validating the method's effectiveness and accuracy [12]. Real-time evaluation requires efficient processing and analysis of large datasets to enable timely data collection and feedback. Data privacy and security pose prominent challenges due to sensitive user data involved. Ensuring effective evaluation while protecting privacy is critical. Limited resources and costs make advanced evaluation methods expensive and resource-intensive, necessitating approaches to conduct effective evaluations under constrained conditions. Interdisciplinary collaboration presents significant challenges in coordinating multidisciplinary evaluations and integrating knowledge from various fields.

7. Future Development Trend of Man-machine System Evaluation

The future of AI and big data-driven human-machine evaluation will witness several significant trends, driving assessment methods toward greater intelligence, precision, and personalization. Numerous scholars have conducted research in this field. Cao Wei et al. highlighted elevators as critical vertical transportation tools, emphasizing the importance of elevator risk assessment. They summarized traditional elevator risk evaluation methods and detailed approaches for collecting and processing large-scale test data, including data collection channels, processes, and data cleaning preprocessing techniques [13]. Multimodal perception and context-aware evaluation leverage AI to integrate multi-modal data such as eye movements, EEG, and physiological signals, providing more comprehensive assessments that enhance evaluation comprehensiveness and accuracy while better understanding user experience [14]. The integration of augmented reality (AR) and virtual reality (VR) creates mixed reality, enabling seamless interaction between virtual and real worlds to deliver more immersive experiences. Hardware advancements include lightweight, high-resolution head-mounted displays, along with more advanced sensors and processors to improve user experience and promote device adoption. These trends will drive the convergence of AR and VR technologies, expand application scenarios, enhance user experience, and spur innovation across industries.

8. Conclusions

Human-machine system evaluation aims to optimize system performance, enhance user experience, and ensure safety. Through multidimensional assessment, it provides insights into current human-machine interaction practices and establishes a foundation for improvement. Quantitative analysis using metrics like task completion time and error rates enables precise measurement of system efficiency and reliability. For instance, in software operation testing, recording users' task completion times and

error counts helps evaluate system usability [15]. Qualitative analysis collects subjective feedback through user interviews and observation methods, observing user behavior to identify potential issues regarding system interfaces and functionalities. By integrating traditional and modern evaluation approaches, challenges such as user diversity, multimodal data integration, real-time assessment, data privacy and security, cost and resource management, and interdisciplinary collaboration are comprehensively evaluated. Despite numerous challenges, the continuous expansion of future development trends has led to advancements in intelligent lighting systems. Researchers like Mo Jusheng have explored the integration of multi-modal data (including eye movements, EEG, and physiological signals) in smart lighting systems, implemented immersive assessments in VR and AR environments, and utilized affective computing and recognition to evaluate user emotional states. These efforts enhance energy efficiency through context-aware and multi-modal interaction technologies, providing users with more natural and intuitive operation methods [10]. Brain-computer interface (BCI) technology is employed to directly assess cognitive and emotional states, while big data analytics help identify behavioral patterns. Technological progress and interdisciplinary collaboration in evaluation methods will continuously optimize human-machine systems, creating efficient, user-friendly, and secure solutions that adapt to evolving user needs and technological advancements. Future research should focus on multi-modal data fusion, AI and machine learning, VR/AR technologies, real-time continuous assessment, and personalized evaluation approaches. Through interdisciplinary collaboration and technological innovation, we can drive intelligent and precise evaluation methods to improve system performance and user experience. Meanwhile, enhanced privacy protection and ethical considerations are essential to ensure compliance and security in the evaluation process.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding

This work was supported by the Regular Project of the Postgraduate Education and Teaching Reform Project of Inner Mongolia Autonomous Region: Research on the Education and Teaching of Intelligent Interaction Design under the Cross-integration of Design and Technology, JG2024055C.

References

- [1] Tian, F.; Gu, T.T. Research on Multi-dimensional Design Optimization of Scientific Research Achievement Management Platform from the Perspective of User Experience. *Packag. Eng.* 2024, 45, 48–56.
- [2] Hao, J.H.; Tian, F.; Hu, Z.K. Review of "Usability Assessment" of Public Self-service Terminal Interface. In Proceedings of the 17th Hunan Graduate Innovation Forum "Better Life, Culture and Tourism Integrated Design Future" Sub-forum; Hunan University of Technology, School of Packaging Design and Art, Ed.; 2024, p. 14.

- [3] Guo, Z.D.; Zhang, Y.X.; Chang, X.Y. Design and Implementation of Intelligent Sleep Aid Instrument Based on Brain Wave Signal. *Electron. Technol.* 2024, 53, 224–225.
- [4] Li, J.; Bu, H.J.; Ming, J.R. Study on the Influencing Factors of User Transfer Use Behavior of AI Literature Reading Tool. *Libr. Constr.* 2025, DOI: 10.13833/j.cnki.isSN1001-8867.20250218.1035.006.
- [5] Wang, S.W.; Zhang, X.Y.; Xie, H.; et al. Word of Mouth Analysis of Xinjiang Fresh Apricot Market Based on Sentiment Analysis of E-commerce Comment Text. *Technol. Ind.* 2025, 25, 143–149.
- [6] Zhang, A.; Ji, C.L.; Bi, X.; et al. Comprehensive Evaluation Method for the Reliability of Complex Systems Considering Operational Profiles. *Electron. Prod. Reliab. Environ. Test.* 2024, 42 (S1), 43-47.
- [7] Tian, C.Z.; Song, M.; Tian, J.W.; et al. Evaluation Method of Human-computer Interaction Efficiency of Command and Control System Based on Eye Movement Tracking. *Aeronaut. Weaponry*, 2023, 30, 57-66.
- [8] Cao, W.; Liu, Y.F.; Xiong, J. Application of Large Sample Test Data to Elevator Risk Assessment. In Proceedings of the 2024 Smart Building and Economic Development Forum under the Vision of New Quality Productivity (III); China Building Metal Structure Magazine Co., LTD, Ed.; 2024, p. 2.
- [9] Huang, W.W.; Chen, L.L. Research on Optimization of Computer Digital Color Extraction and Collocation Technology Based on Visual Perception. *Autom. Instrum.* 2025, 197-200+205.
- [10] Mo, J.S.; Wei, S.L.; Zhao, L.H.; et al. Context Perception and Multi-modal Interaction Technology in an Intelligent Lighting System. *China Light. Appl.* 2024, 134-137.
- [11] Ma, Z.Q.; Cui, X. Human-computer Collaborative Teaching Model from the Perspective of Intersubjectivity: Theoretical Framework and Practical Form. *E-educ. Res.* 2026, 3, 14-21.
- [12] Ni, Y.; Luo, M.; Zhao, D.H. Research and Experiment on Enhanced Interaction Strategy of Autonomous Driving Based on Exterior Human-machine Interface eHMI. *China J. Highw. Transp.* 2026, 1-22.
- [13] Shi, S.Y.; Peng, X.X.; Zhang, Y.H.; Zhao, F.F.; Ye, X.S. Research on the Mechanism of Human-machine Interaction on Project Resilience under the Background of Industry 5.0: From the Perspective of Knowledge and Ability. *Technol. Econ.* 2026, 45, 91-101.
- [14] Wang, Y.W.; Zhao, X. Review of Augmented Reality Human-machine Interaction Technology from the Dual Perspectives of Input and Output. *J. China Image Graph.* 2026, 31, 349-373.
- [15] Shi, W.G.; Wang, W.Q.; Wang, J.X.; Liu, S.D.; Liang, X.; Hou, Z.G. Active Adaptive Human-machine Interaction Control for Robot-assisted Gait Training. *Robot.* 2026, 1-9.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)