HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

DEVELOPMENT AND EVALUATION OF A SIMULATION-BASED TRAINING SYSTEM FOR CREW PREPARATION ON THE ZSU-23-4 "SHILKA" SELF-PROPELLED ANTI-AIRCRAFT SYSTEM

Nurmatov Boburbek Mirkhojiddinovich

Lieutenant Colonel, Senior Lecturer, Department of Ground Forces Air Defense

E-mail: boburnurmatov1987@gmail.com

Doi: https://doi.org/10.5281/zenodo.24532369

Abstract. This article examines the scientific and methodological foundations for developing and evaluating a simulation-based training system for the crew of the ZSU-23-4 Shilka self-propelled anti-aircraft system. It discusses the educational potential of simulation technologies, the functional structure of the proposed simulator, and criteria for assessing training outcomes.

Keywords: ZSU-23-4 Shilka, simulation-based training, military crew, air defense, virtual learning, professional training, effectiveness evaluation.

Annotatsiya. Mazkur maqolada ZSU-23-4 «Shilka» o‘ziyurar zenit qurilmasi jangovar hisobi mutaxassislarini tayyorlash uchun simulyatsion trenajyorni ishlab chiqish va uning samaradorligini baholashning ilmiy-metodik asoslari yoritilgan. Simulyatsion texnologiyalarning o‘quv jarayonidagi imkoniyatlari, trenajyorning funksional tuzilishi hamda tayyorgarlik natijalarini baholash mezonlari ko‘rib chiqilgan.

Kalit so‘zlar: ZSU-23-4 «Shilka», simulyatsion trenajyor, jangovar hisob, havo hujumidan mudofaa, virtual ta’lim, kasbiy tayyorgarlik, samaradorlikni baholash.

Аннотация. В статье рассматриваются научно-методические основы разработки и оценки эффективности симуляционного тренажёра для подготовки боевого расчёта самоходной зенитной установки ЗСУ-23-4 «Шилка». Анализируются возможности симуляционных технологий в учебном процессе, функциональная структура тренажёра и критерии оценки результатов подготовки специалистов.

Ключевые слова: ЗСУ-23-4 «Шилка», симуляционный тренажёр, боевой расчёт, противовоздушная оборона, виртуальное обучение, профессиональная подготовка, оценка эффективности.

Introduction

Training personnel to work collaboratively with complex technical systems requires more than the acquisition of theoretical knowledge. Learners must also develop a shared understanding of responsibilities, communicate consistently, recognize procedural errors, and maintain performance under changing instructional conditions. Conventional teaching methods, including lectures, demonstrations, and supervised practice, remain valuable; however, they may provide limited opportunities for repeated practice, individualized feedback, and standardized assessment. Simulation-

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

based education addresses some of these constraints by allowing instructors to reproduce learning situations, vary task complexity, and document learner responses in a controlled environment.

The ZSU-23-4 "Shilka" is a historically established self-propelled anti-aircraft platform. In this article, the platform serves as the contextual reference for a proposed crew-training environment. The study deliberately addresses pedagogical organization, non-operational interface familiarization, teamwork, communication, and educational assessment. It does not specify weapons employment, engagement procedures, target-acquisition logic, fire-control algorithms, tactical optimization, or operational performance parameters.

Evidence from the broader simulation-training literature suggests that realism alone does not determine instructional effectiveness. A meta-analysis by Doozandeh and Hedayati found that the relationship between simulation fidelity and transfer of learning depends on factors including prior expertise and the characteristics of the system being taught. Consequently, a simulator should be designed around measurable learning objectives rather than the assumption that maximum visual or technical fidelity is always preferable [1].

A second methodological challenge concerns evaluation. A visually convincing simulator may still be difficult to use, fail to provide meaningful feedback, or produce little improvement in transferable knowledge. Software product quality and educational effectiveness should therefore be examined as related but distinct constructs. ISO/IEC 25010:2023 provides a general product-quality framework for software and information and communication technology products [2]. The four-level training evaluation model distinguishes learners' reactions, learning outcomes, behavior or transfer, and broader results [3].

The aim of this study is to develop a conceptually structured simulation-based training system for the educational preparation of ZSU-23-4 crew members and to propose a scientifically defensible procedure for evaluating its quality and potential instructional effectiveness.

Materials and Methods

A design-oriented methodological study was selected. Its purpose was to specify a training-system architecture and an evaluation protocol rather than to report the outcomes of an implemented experimental intervention. The approach comprised four sequential stages: analysis of educational needs; formulation of instructional and software requirements; development of a modular conceptual design; and specification of validation and evaluation procedures. These stages form a proposed development pathway and should not be interpreted as evidence that a working prototype or field trial has already been completed.

The target educational outcomes were defined at the level of general professional preparation: understanding the structure of a simulated work environment; recognizing assigned crew roles; exchanging clear information within a team; completing instructor-defined, non-operational procedural learning tasks; detecting deviations from an instructional checklist; and reflecting on performance during structured debriefing. Learning activities were designed to avoid representation of live engagement decisions or system-specific operational functions.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

The proposed requirements were divided into pedagogical, interaction, technical, and assessment categories. Pedagogical requirements include explicit learning objectives, progressive difficulty, practice repetition, and formative feedback. Interaction requirements include readable displays, consistent controls, accessible instructions, and role clarity. Technical requirements include stable operation, controlled user access, and reliable record keeping. Assessment requirements include transparent scoring criteria, reproducible scenarios, and exportable instructor reports.

The proposed system contains five modules. The instructor module is intended to assign learning objectives, configure generic instructional situations, observe participation, and initiate debriefing. The trainee module provides a virtual interface for orientation, task rehearsal, and role-based collaboration. The scenario module stores non-operational educational exercises and supports repetition under comparable conditions. The assessment module records completion status, communication quality, checklist adherence, and requests for assistance. The feedback module presents a chronological summary of decisions, instructor comments, and opportunities for improvement.

For implementation, modules should be separated through clearly documented interfaces so that educational content can be revised without changing the entire application. Event logs should contain only information necessary for training assessment. Data governance should define authorized access, retention periods, and secure handling of trainee records. Any eventual integration with real equipment would require an independent safety, cybersecurity, and institutional authorization review and lies outside the present study.

Table 1. Proposed modules and educational functions

Module	Principal educational function	Proposed evaluation indicator
Instructor interface	Assignment, supervision, and feedback	Instructor-rated clarity and workflow completion
Trainee interface	Orientation and guided rehearsal	Task completion and perceived usability
Scenario library	Repeatable, non-operational learning situations	Scenario consistency and objective coverage
Assessment engine	Documentation of learning performance	Scoring agreement and completeness of records
Debriefing module	Structured reflection and improvement	Feedback usefulness and corrective learning

Table 1 demonstrates that the architecture is organized around educational functions rather than the replication of weapon-control processes. Each module is paired with a measurable indicator, enabling a future prototype to be evaluated systematically. The modular structure may also facilitate adaptation to different learner experience levels and instructional settings.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

Educational scenarios should be constructed from explicit learning outcomes and a standardized template containing the scenario title, prerequisites, participant roles, instructor prompts, expected communication behaviors, completion criteria, and debriefing questions. An introductory scenario may emphasize interface orientation and role identification; an intermediate scenario may focus on the coordination of generic checklist tasks; and an advanced educational scenario may involve managing competing non-operational administrative demands while maintaining clear communication. Scenario difficulty should be adjusted by instructional complexity and information load, not by optimizing weapon performance.

A structured debriefing is proposed after each exercise. The instructor should invite trainees to describe their actions, identify communication problems, compare their performance with predefined criteria, and formulate specific improvement goals. This approach creates a feedback loop connecting simulated experience with reflection and subsequent practice.

A future empirical study could employ a prospective, controlled pretest–posttest design. Eligible participants would be adult trainees enrolled in an authorized educational program. The intervention group would receive the conventional curriculum supplemented by the proposed simulator, while the comparison group would receive the conventional curriculum without the simulator during the study period. Allocation procedures, sample size, eligibility criteria, and intervention duration should be registered in a study protocol before data collection. If random allocation is feasible, it should be preferred; otherwise, baseline differences should be documented and addressed analytically.

The primary educational outcome should be a standardized assessment of non-operational procedural knowledge and crew communication using a common rubric administered before and after training. Secondary outcomes may include retention at a delayed follow-up, instructor-rated collaboration, perceived workload, satisfaction, and usability. Assessors should be blinded to group assignment where practical. The scoring rubric should be piloted, and inter-rater agreement should be examined before primary analysis.

Software quality assessment may be guided by relevant dimensions of ISO/IEC 25010:2023, including functional suitability, performance efficiency, interaction capability, reliability, security, maintainability, flexibility, safety, and compatibility as appropriate to the final product specification [2]. The four-level training evaluation framework may be used to distinguish immediate satisfaction from demonstrated learning and subsequent transfer [3].

Table 2. Proposed evaluation matrix

Evaluation domain	Indicator	Data source	Interpretation
Learning	Change in standardized knowledge score	Pretest and posttest	Educational improvement, not proof of transfer

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

Teamwork	Communication and role-coordination rubric	Independent observer ratings	Quality of collaborative behavior
Usability	System Usability Scale or validated alternative	Trainee questionnaire	Perceived ease of use
Reliability	Rate of completed sessions without software failure	System event log	Technical dependability
Assessment quality	Inter-rater agreement	Parallel scoring of sample sessions	Consistency of ratings
Retention	Delayed assessment score	Follow-up assessment	Persistence of acquired knowledge

Table 2 separates learning indicators from technical and user-experience indicators. This distinction is essential because favorable usability ratings do not necessarily demonstrate learning gains. Similarly, short-term test improvement does not establish transfer to authentic professional settings without additional evidence.

For a future controlled study, descriptive statistics should summarize baseline characteristics, test scores, questionnaire results, and missing data. The primary analysis could estimate the difference in posttest performance between groups while adjusting for baseline score through an appropriate regression or analysis-of-covariance model. Effect sizes and 95% confidence intervals should be reported alongside statistical significance. If assumptions are not met, robust or nonparametric alternatives should be considered. Repeated measurements and clustering by training group or instructor should be accounted for when relevant.

Sample size should be justified through an a priori power calculation based on the primary outcome and a plausible minimally important difference. Missing-data procedures and the primary analysis plan should be specified before outcomes are examined. Any exploratory subgroup analyses should be identified as such. Ethical approval, institutional authorization, voluntary informed consent, confidentiality protections, and secure storage of trainee data would be required before participant recruitment. The proposed research should not collect or publish restricted operational information.

Results

The principal result of this design-oriented study is a structured conceptual model for an educational simulator. The model links instructional objectives to scenario design, learner interaction, performance recording, and debriefing. It specifies the responsibilities of the instructor and trainee interfaces and provides a basis for producing testable software requirements. Because no operational prototype was built or evaluated within the scope of the present article, these results describe design outputs rather than experimentally observed training outcomes.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

A second result is the specification of a progression model with three instructional levels. The introductory level concentrates on orientation and the understanding of role responsibilities. The intermediate level emphasizes accurate execution of generic learning checklists and clear communication. The advanced level combines multiple non-operational educational tasks, requiring trainees to maintain coordination and document their decisions. Progression between levels should be based on predefined mastery criteria rather than time spent in the simulation alone.

The proposed evaluation matrix provides a direct correspondence between research questions, indicators, and sources of evidence. Content validity can be assessed by a panel of educational-methodology, human-factors, and software-quality specialists. The expert panel should independently rate the relevance and clarity of learning objectives, tasks, interface elements, and assessment criteria. Disagreements should be documented and used to revise the design. A subsequent pilot study should examine feasibility, instructions, recording accuracy, and the time required to complete educational scenarios.

The evaluation protocol also identifies important safeguards against overinterpretation. It distinguishes subjective satisfaction from objective learning, recommends baseline adjustment, requires reporting of uncertainty, and includes a delayed assessment where feasible. These provisions support more credible conclusions than a simple post-training satisfaction survey or an uncontrolled before-and-after comparison.

Table 3. Development stages and required evidence

Stage	Expected output	Evidence required before advancement
Requirements analysis	Approved learning objectives and user requirements	Expert confirmation of relevance and scope
Prototype development	Functional educational modules	Documented software verification
Pilot testing	Revised interface and scenarios	Feasibility and usability evidence
Controlled evaluation	Comparative learning outcomes	Predefined analysis with uncertainty estimates
Implementation review	Decision on instructional adoption	Independent review of benefits and limitations

Table 3 outlines a staged validation process in which the quality of evidence increases progressively. A conceptually complete architecture does not justify claims about learning effectiveness until a functioning system has undergone empirical evaluation. This staged approach reduces the risk of adopting an educational technology on the basis of appearance or perceived novelty alone.

No participant sample, trial dataset, usability scores, software reliability measurements, or comparative effectiveness estimates were available for this article. Accordingly, no numerical results,

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

p-values, confidence intervals, or improvement percentages are reported. The absence of empirical observations is a limitation of the present design study, but explicitly acknowledging this limitation preserves the distinction between a proposed methodology and validated findings.

Discussion

The proposed training-system framework responds to a general educational need: complex team-based learning benefits from repeatable exercises, structured observation, and timely feedback. Its central design choice is to prioritize instructional fidelity—how well the simulation represents the learning task—over unrestricted replication of technical system behavior. This decision is consistent with evidence that higher simulation fidelity does not always produce better transfer and that the optimal level depends on learner characteristics and training objectives [1].

The model also emphasizes the distinction between technical quality and educational effectiveness. A stable and accessible interface may improve learners' experience, yet instructional value ultimately depends on what trainees learn and retain. The ISO/IEC 25010:2023 framework can support systematic product-quality assessment, while a training evaluation framework can help organize evidence across satisfaction, learning, behavior, and broader results [2, 3]. These frameworks are complementary rather than interchangeable.

A further strength of the proposed approach is the use of explicit assessment criteria. Transparent rubrics can improve consistency between instructors and help trainees understand expectations. Nevertheless, the reliability and validity of these rubrics must be demonstrated empirically. Expert agreement alone is insufficient to establish that simulator-based performance predicts future professional competence.

Several limitations should be noted. First, the article presents a conceptual design rather than a deployed simulator. Second, no empirical data are available to estimate usability, learning gains, retention, or transfer. Third, the educational scenarios intentionally exclude operational weapons functions, so the framework cannot support claims about combat readiness or technical proficiency with live equipment. Fourth, the usefulness of the proposed system may vary with instructor preparation, learner experience, institutional resources, and local curriculum requirements.

Future research should therefore proceed in stages: expert review, prototype development, formative usability testing, pilot feasibility assessment, and a preregistered comparative educational study. Longer-term research could investigate whether improvements in generic communication, coordination, and procedural learning persist across different instructional settings. Such work should use independent outcome assessment and comply with all relevant institutional, ethical, and information-security requirements.

Conclusion

This article presents a conceptual framework for developing and evaluating a simulation-based educational system associated with crew preparation for the ZSU-23-4 "Shilka" platform. The framework combines modular instructional architecture, role-based non-operational exercises, standardized assessment, structured debriefing, and a prospective evaluation protocol. Its contribution

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

is methodological: it identifies how educational objectives can be translated into simulator requirements and how future research can distinguish usability, learning, retention, and transfer. The proposed approach has not yet been validated through implementation or participant testing. Accordingly, claims of educational effectiveness must await controlled empirical evaluation.

References

1. Doozandeh, P., & Hedayati, S. (2022). The effect of simulation fidelity on transfer of training for troubleshooting professionals: A meta-analysis. IISE Transactions on Occupational Ergonomics and Human Factors, 10(3), 117–134. https://doi.org/10.1080/24725838.2022.2099483
2. International Organization for Standardization & International Electrotechnical Commission. (2023). ISO/IEC 25010:2023: Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—Product quality model. https://www.iso.org/standard/78176.html
3. U.S. Office of Personnel Management. (2011). Training Evaluation Field Guide. https://www.opm.gov/policy-data-oversight/training-and-development/reference-materials/training_evaluation.pdf
4. Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). Evaluating Training Programs: The Four Levels (3rd ed.). Berrett-Koehler Publishers.
5. Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), Usability Evaluation in Industry (pp. 189–194). Taylor & Francis.
6. Salas, E., Wildman, J. L., & Piccolo, R. F. (2009). Using simulation-based training to enhance management education. Academy of Management Learning & Education, 8(4), 559–573. https://doi.org/10.5465/amle.8.4.zqr559