p-ISSN 1607-3274 Радіоелектроніка, інформатика, управління. 2026. № 2
e-ISSN 2313-688X Radio Electronics, Computer Science, Control. 2026. № 2

UDC 004.9

INFORMATION TECHNOLOGY FOR ANALYSING AND QUANTIFYING THE EFFECTIVENESS OF VR TRAINING FOR FIRST AID SKILLS IMPROVEMENT IN EMERGENCIES BASED ON BEHAVIOURAL AND STATISTICAL MODELS

Vysotska V. – PhD, Professor of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine. ROR: https://ror.org/0542q3127. ORCID: https://orcid.org/0000-0001-6417-3689.

Chyrun S. – Student of the Information and Communication Technologies Department, Lviv Polytechnic National University, Lviv, Ukraine. ROR: https://ror.org/0542q3127. ORCID: https://orcid.org/0000-0002-2829-0164.

ABSTRACT

Context. The study's relevance stems from the urgent need to improve the effectiveness of prehospital first aid training under heightened risk to civilian populations, particularly in emergency scenarios involving damage to civilian infrastructure. Traditional training approaches are limited in their ability to realistically simulate hazardous situations, objectively monitor participants' actions, and quantitatively analyse learning dynamics. Virtual reality (VR) technologies enable the creation of fully controlled and repeatable simulation environments with automated logging of temporal, behavioural, and performance-related parameters, providing new opportunities for objective assessment of training effectiveness.

Objective of the study is to develop and experimentally validate an information technology for the quantitative evaluation of VR-based training effectiveness in developing prehospital first aid skills, compared with traditional training methods.

Method. An experimental study was conducted using a controlled design with VR and control groups, including pre-test, post-test, and delayed retention measurements. Training effectiveness was evaluated using a set of quantitative metrics, including reaction time RT, action accuracy, number of critical errors, Precision, Recall, F1-score, and a composite performance score S. Learning dynamics were analysed using exponential learning curve models, mixed-effects models for repeated measurements, parametric and non-parametric statistical tests, bootstrap confidence intervals, and effect size estimation (Cohen's d).

Results demonstrate a statistically confirmed advantage of VR-based training over traditional methods. The average reaction time for critical actions in the VR group was reduced by approximately 10–20% compared to the control group (e.g., ≈34 seconds vs. ≈40 seconds in bleeding control scenarios). Action accuracy increased from approximately 0.78 in the control group to 0.86 in the VR group, corresponding to an improvement of about 8–10%. The composite performance score S was higher in the VR group by 0.05–0.12 (on a 0–1 scale), depending on the scenario. F1-scores for automated action classification reached 0.90–0.92, and large effect sizes were observed, with Cohen's d values up to approximately 2.3. Retention testing further indicated improved stability and long-term preservation of skills following VR-based training.

Conclusions. The proposed information technology and experimental results support the use of VR as an effective, scalable, and data-driven approach for prehospital first aid training for civilians, emergency responders, and medical personnel in emergency and disaster-response contexts.

KEYWORDS: virtual reality, prehospital first aid, VR-based training, simulation-based learning, training effectiveness evaluation, composite performance score, learning curves, mixed-effects modelling, skill retention, emergency scenarios.

ABBREVIATIONS

GAI – generative artificial intelligence;
AI – Artificial Intelligence;
AR – Augmented Reality;
VR – Virtual Reality;
WBS – Work Breakdown Structure;
MVP – Minimum Viable Product;
STAI – State-Trait Anxiety Inventory;
SAM – Self-Assessment Manikin.

NOMENCLATURE

i – index of the participant (i = 1.2,...,N);
g – training group ($g \in \{0,1\}$), where g=1 for VR, g =0 – Control;
RT – reaction time/ operation execution time (s);
A – accuracy of actions/operations (binary or partial assessment, 0–1);
Err – number of critical errors;
Prs – level of immersion (presence) (self-esteem, scale 1–7);
Str – stress level (surveys or biometric data, for example, heart rate);
S – composite performance score / composite score for the scenario;
w_j – metric weights (sum 1);
$\bullet'$ – normalisation to the range [0,1];
S_0 – initial skill level (pre-test, at the start of training);
k – coefficient of learning speed (the higher, the faster it is absorbed);
$S_{\max}$ – theoretically possible maximum (learning asymptote, asymptotic skill level);
u_i – random effect of the participant;
β_3 – interaction parameter (Group × Session).
$P_i(t)$ – indicator of the participant's productivity/skill depending on the training time t;
P_0 – beginner level;
P_∞ – asymptotic maximum;
a,b,c – customised parameters;
Δ – expected difference in averages (effect);
σ^2 – expected variance of the S indicator;

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5

OPEN ACCESS

z – critical values of the normal distribution for the level of significance of α and power 1−β;

λ – loss rate;

P_{post} – level immediately after training;

$mean_{VR}$ / $mean_C$ – average values of the integral efficiency indicator S for the VR group and the control group;

t-pval, MW-pval – two-sample tests (t-test and Mann-Whitney). Values < 0.05 indicate statistically significant differences;

Cohen's d – effect size > 0.8 – large;

p(Group) – main effect of the group from an extended mixed model with a random slope for the Session; p(Group×Session) – interaction (difference in learning speed between VR and control);

Δ_{boot} ($VR–C$) і 95% CI – bootstrap estimate of the average difference between VR and Control; if the interval does not cover 0, the effect is stable.

$S(t)$ – level of skill formation at time t;

t – measurement stages (pre, Post, retention), $t=\{0,1,2\}$, session number or training time (in hours, weeks, etc.);

S_{it} – indicator of the effectiveness of the i-th participant in session T;

$Group_i$ – binary variable (0 – control, 1 – VR);

$u_i \sim N(0,\sigma_u^2)$ – random effect (individual tilt/intercept);

$\epsilon_{it} \sim N(0,\sigma_\epsilon^2)$ – residual error;

u_{1i} – individual differences in the pace of learning (speed).

μ_{VR} – average value of the integral efficiency indicator S in the VR group;

μ_{CTRL} – average for the control group;

$\dot{X}_i$ – average values in groups;

s_i^2 – sample variances;

n_i – sample sizes;

R_1 – sum of the ranks of the first group;

Φ – normal distribution function;

$z_{1-\alpha/2}$ – quantile of the standard normal for the bilateral α (for α=0.05 it is 1.96),

$z_{1-\beta}$ – quantile for power (for power 0.80, it is 0.84);

m_c – average cluster size (number of participants in one cluster);

ρ_{ICC} – intra-cluster correlation (ICC);

r_{ij} – Participant I's response to item j;

r_{PS} – correlation coefficient between presence and skill performance;

S_i – integral effectiveness of training for the first participant;

β_P – immersion effect;

β_R – effect of stress;

β_{PR} – interaction of presence × stress;

$r_{PS \cdot R}$ – partial correlation between P and S, controlled by R.

INTRODUCTION

The development of effective methods of teaching pre-medical care is critically important in today's conditions of increased risks for the civilian population, in particular as a result of armed conflicts and damage to civilian infrastructure. Traditional approaches to the training of specialists and volunteers are often limited by the ability to model dangerous scenarios, monitor participants' behavioural parameters, and develop automated skills in realistic conditions. Virtual reality (VR) enables the creation of fully controlled, playable, and highly dynamic simulations that provide deeper immersion and allow you to accurately capture the quality and timing of actions. Experimental data demonstrate a consistent advantage of VR training in most scenarios, including bleeding control, CPR, and triage, further emphasising the relevance of this study.

Despite the significant potential of VR technologies, quantifying their effectiveness in first aid during emergencies remains understudied. Existing skills assessment methods often do not provide comprehensive data on the accuracy of actions, response speed, learning dynamics, and the interaction between immersion and stress factors. The lack of standardised analytical models and mathematical descriptions of the learning process is also a problem, making it difficult to objectively compare VR and traditional methods. Therefore, there is a need for a comprehensive experimental study using validated metrics, mixed statistical models, and a clear assessment structure.

The purpose of the study is to develop an information technology for quantitative assessment of the effectiveness of VR training in the formation of pre-medical care skills and to determine the advantages of this technique over traditional training. The study is based on experimental substantiation of the effectiveness of VR training in developing first-aid skills under simulated damage to civilian infrastructure, as well as on the analysis of the parameters of educational dynamics using modern statistical and mathematical models.

To achieve this goal, it is necessary:

– Develop an experimental scheme for assessing skills in VR and the control group, including pre-/post-/retention measurements.

– Determine key performance indicators (reaction time, accuracy, F1-score, composite score S) and ensure their correct measurement.

– Build mathematical models of learning curves and mixed effects models to analyse the dynamics of skill acquisition.

– Conduct a comparative analysis between VR and traditional learning methods using t-tests, Mann-Whitney tests, bootstrap assessments, and Group×Session models.

– To analyse the impact of immersion (presence) and stress on the effectiveness of actions.

– Evaluate statistical significance, effect size (Cohen's d), and stability of results through confidence intervals and bootstrapping.

The object of research is the process of developing and evaluating practical first-aid skills in simulated dangerous situations. The study examines methods, models, and metrics for assessing the effectiveness of VR

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
CC BY SA

training relative to traditional first-aid training, including learning dynamics, behavioural indicators, and statistical characteristics of actions. For the first time, the study integrates a comprehensive system of quantitative metrics, automated logs of the VR environment, and mixed mathematical models to assess the quality of pre-medical skills in dynamics. The novelty lies in combining exponential learning models, mixed-effects models, and bootstrap assessment to compare VR and traditional training, thereby providing a highly accurate evaluation of group and temporal effects. The data obtained demonstrate significant effect sizes (up to $d≈2.3$) and a statistically confirmed advantage of VR in the formation of key skills.

The results of the study can be used to create effective pre-medical care training programs for civilians, military, rescuers and medical personnel. The proposed approaches allow optimising educational processes, introducing standardised assessment models and improving the quality of training in realistic but safe simulated conditions. VR trainings have the potential to be scalable, affordable, and cost-effective, which significantly expands their application in security, education, and crisis response systems.

1 ROBLEM STATEMENT

The goal is to formalise a mathematical model to evaluate the effectiveness of VR training in developing first-aid skills and to determine statistically significant differences between the VR and control groups using quantitative metrics, training models, and statistical analysis methods. For each participant i (i=1,2,..., N), a measurement vector is observed:

$$x_{i,t}=(RT_{i,t}, A_{i,t}, Err_{i,t}, Prs_{i,t}, Str_{i,t}, S_{i,t}). \quad (1)$$

The composite score is formed as:

$$S_{i,t}=w_1A'_{i,t}+w_2RT'_{i,t}+w_3(1-Err'_{i,t})+w_4Prs'_{i,t}. \quad (2)$$

It is necessary to assess whether VR training provides a significant increase in efficiency:

$$H_0:\mu_{VR}=\mu_C,\quad H_1:\mu_{VR}>\mu_C, \quad (3)$$
$$\mu_{VR}=E[S_{i,t}^{(VR)}],\ \mu_C=E[S_{i,t}^{(C)}]. \quad (4)$$

The exponential model describes the process of skill formation:

$$S_i(t)=S_{\max}-(S_{\max}-S_0)e^{-kt}. \quad (5)$$

For the two groups, the models look like:

$$S^{VR}(t)=S^{VR}_{\max}-(S^{VR}_{\max}-S_0^{VR})e^{-k_{VR}t}, \quad (6)$$
$$S^C(t)=S^C_{\max}-(S^C_{\max}-S_0^C)\,e^{-k_Ct}. \quad (7)$$

The task is to evaluate the parameters

$$\theta=\{k_{VR}, k_C, S^{VR}_{\max}, S^C_{\max}\}. \quad (8)$$

and check whether $k_{VR} > k_C$.

Taking into account individual differences, the mixed-effects model is used:

$$S_{i,t}=\beta_0+\beta_1g_i+\beta_2t+\beta_3(g_i\cdot t)+u_i+\varepsilon_{i,t}. \quad (9)$$
$$u_i\sim N(0,\sigma_u^2),\quad \varepsilon_{i,t}\sim N(0,\sigma^2). \quad (10)$$

The task is to check H_0:β_1=0 and H_0:β_3=0.

The data contains significant effects (up to d=2.3). Formally:

$$d=(\hat{S}_{VR}-\hat{S}_C)/\sqrt{(1/2(\sigma^2_{VR}+\sigma^2_C))}. \quad (11)$$

The task is to calculate d for each scenario and assess whether the effect is significant ($d > 0.8$ indicates a large effect). To evaluate the stability of the mean difference:

$$\Delta=\hat{S}_{VR}-\hat{S}_C \quad (12)$$

Multiple bootstrap samples are built:

$$\Delta^{(b)}=\hat{S}_{VR}^{(b)}-\hat{S}_C^{(b)},\ b=1,\ldots,B. \quad (13)$$

95% confidence interval:

$$CI_{95\%}=[\Delta^{(2.5\%)},\ \Delta^{(97.5\%)}]. \quad (14)$$

The task is to check whether $0 \notin CI_{95\%}$. Find parameters:

$$\Theta=\{\mu_{VR}, \mu_C, k_{VR}, k_C, \beta_1, \beta_3\}, \quad (15)$$

such that:

– VR has a higher average skill level $\mu_{VR} > \mu_C$;
– Learns faster $k_{VR} > k_C$;
– Demonstrates a significant group effect $\beta_1>0$;
– Has accelerated learning in time: $\beta_3>0$;
– The difference between the groups is stable $0 \notin CI_{95\%}$.

It constitutes a formally defined scientific task of the study.

2 REVIEW OF THE LITERATURE

Over the past decade, the application of virtual reality (VR) technologies in medical education and pre-medical/rescue training has increased significantly. Such approaches promise to provide a safe, repeatable, and controlled environment for practising clinical or emergency skills, which is often impossible or costly in real-world training settings [1–3].

A systematic review encompassing more than 50 studies demonstrated the multifaceted application of VR, ranging from surgical simulators to training for emergency scenarios. The authors note that VR improves learning quality, enhances interactivity, and enhances the effectiveness of distance or blended learning, especially during a pandemic or with limited resources [1, 4]. A meta-analysis of the impact of VR training in medical and nursing education found a statistically significant improvement in skill score compared to traditional methods: mean standard deviation (SMD) of

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
CC BY SA

approximately 0.23 [0.11, 0.34] for general skills, and SMD ≈0.32 for medical and nursing disciplines. At the same time, VR training also reduced task completion time – SMD ≈0.59 for general skills and 0.74 for medical skills [2].

In a study comparing VR training and a dummy simulation for basic rescue skills (CPR, AED, first aid, "Stop the Bleed"), the VR training group performed significantly better in most scenarios (AED, choking management, CPR), while in the "Stop the Bleed" scenario, the difference was not statistically significant [3]. It demonstrates that VR can be especially effective for procedures where precise algorithms are critical – chest compression, AED application, and basic skills that require repetition and automatism.

In the context of Basic Life Support (BLS), a study in the student group showed that after the VR course, the level of subjective learning gain was higher, and the "no-flow time" during simulated cardiopulmonary resuscitation (CPR) was no worse than that of the classic course [5–6]. Also, a small study among volunteers (non-health workers) showed that the VR group retained knowledge and theoretical training better for 12 months than the traditional simulation group [7]. Thus, VR is exceptionally well-suited for scaling up basic life skills.

Along with positive conclusions, the literature also contains critical comments. For example, in one of the pilot studies on the use of a tourniquet for bleeding control in EMT students, the addition of VR training did not lead to a statistically significant improvement in the success of tourniquet placement 70 days after training: 63% in the control versus 57% in the VR group ($p = 0.57$) [8]. The main reasons for failures in the VR group were insufficiently realistic haptics (insufficient tightening) and failure to follow the procedure steps, as in the control group [8]. A similar study among emergency medicine residents showed that 3 months after the training, the level of success in using the tourniquet (according to the standards) in the VR group and in the group with instructors was approximately the same (92–97% on Day 0; 90–95% on retention), without statistically significant differences [9]. Thus, VR augmentation does not always guarantee an advantage over the traditional approach – the results may depend on the type of skill, the quality of the simulation, and the availability of feedback. In addition, [10–15] notes the following common problems in VR education: high equipment costs, technical failures, insufficient haptic feedback, potential cognitive overload from high sensory stimulation, and the risk of reduced interpersonal communication if live learning is replaced by VR [1–2]. According to the totality of studies, the following patterns can be distinguished [15–23]:

– VR trainings are most effective for biopsychomotor skills that have a precise sequence of actions and do not require complex tactile feedback (for example, CPR, AED, basic manipulations, procedures with precise algorithms).

– Where the role of tactile feedback (graphic haptics, feeling of pressure, elasticity, etc.) is crucial – such as the use of a tourniquet in a real bleeding situation – VR without additional dummies or haptic devices often shows limited effectiveness.

– VR provides a significant advantage in scalable, repeatable, asynchronous learning, especially when there are no resources for regular live classes or when access to instructors is limited.

– The hybrid approach provides the most benefit: a combination of VR, a dummy/physical simulation and a human instructor, especially for advanced/critical skills.

Taking into account the described advantages and limitations of VR training and, based on the set goal, to assess the effectiveness of VR in the development of first aid skills in conditions of damage to civilian infrastructure, the literature sources confirm the validity of this approach. However, they also indicate the need to:

– ensure high-quality simulation, especially if haptics are needed;

– combine VR with physical mannequins or other methods when it comes to complex manipulations;

– conduct long-term re-testing with retention to assess the retention of skills;

– carefully design the research (randomisation, blind assessment, clear metrics, standardisation of the assessment).

Therefore, our approach – using quantitative metrics, mixed-effects models, retention tests, and simulation scenarios – is consistent with best practices and recommendations from the literature.

3 MATERIALS AND METHODS

Here is a concise, technically accurate description of the research process (methodology, quantitative metrics, and statistical models) for a project on the use of virtual reality (VR) to develop first-aid skills in the face of damage to civilian infrastructure.

Structure of the study (experimental scheme):

1. Definition of goals and hypotheses:

– Main hypothesis H_1 – training in VR improves practical skills (speed and quality of actions) compared to a standard training course.

– Zero H_0 – there are no differences.

2. Design – Randomised Controlled Trial (RCT) or Quasi-Experiment:

– Groups – G_{VR} (receive VR training), G_C (control – traditional training).

– Before/after each participant passes a pre-test and post-test; an additional post-test is possible after T_{ret} weeks to assess the retention of skills. The project provides alpha/beta testing with specialists and volunteers. Let's denote the participant's index i and the training/testing time t. Metrics (calculations): Accuracy (as a proportion of correct actions), Precision / Recall / F1 – if the evaluation of actions to detect/correct critical errors (TP, FP, FN), Composite score (weighted sum of indicators).

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS CC BY SA

$$Accuracy=(\sum_{i=1}^{N} correct_{actionsi})/(\sum_{i=1}^{N} total_{actionsi}), \quad (16)$$

$$Precision=TP/(TP+FP), \quad (17)$$

$$\text{Recall}=TP/(TP+FN), \quad (18)$$

$$F1=2\cdot(Precision\cdot Recall)/(Precision+Recall). \quad (19)$$

The approach is useful in classifying "correctly/incorrectly applied tourniquet", "correct CPR technique", etc.

$$S_i=w_1\cdot(1-RT_i/RT_{\max})+w_2\cdot A_i+w_3\cdot Prs_i+ + w_4\cdot(1-Err_i/Err_{\max}). \quad (20)$$

Learning is often described by the exponential approximation function (asymptotic improvement) or the law of degree (power law of practice):

$$P(t)=P_\infty-(P_\infty-P_0)e^{-kt}, \quad (21)$$

$$P(t)=a\,t^{-b}+c, \quad (22)$$

where $b>0$ reflects the rate of improvement. These models allow you to compare the learning rates k and b between G_{VR} and G_C. Statistically test the difference in parameters (for example, through nonlinear regression and comparison of confidence intervals).

Statistical tests and effect evaluations:

– Level test in pre/post for "before/after" measurements – paired t-test (if normal), or Wilcoxon signed-rank for abnormal distributions:

$$H_0:\mu_{post}-\mu_{pre}=0, \quad (23)$$

– Comparison of the two groups – independent t-test or Mann–Whitney:

$$t=\dot{X}_{VR}-\dot{X}_C/\sqrt{(s^2_{VR}/n_{VR}+s^2_C/n_C)}, \quad (24)$$

– ANOVA / Mixed-effects model – if there are repeated measurements by time and nesting (participants in groups), the model is useful:

$$Y_{it}=\beta_0+\beta_1\cdot\text{Group}_i+\beta_2\cdot t+\beta_3\cdot(\text{Group}_i\cdot t)+u_i+\varepsilon_{it}, \quad (25)$$

– Effect size estimation (Cohen's d):

$$d=(\dot{X}_{VR}-\dot{X}_C)/s_{pooled}, \quad (26)$$

$$s_{pooled}=\sqrt{((n_{VR}-1)s^2_{VR}+(n_C-1)s^2_C)/(n_{VR}+n_C-2)}, \quad (27)$$

where d=0.2 is small, 0.5 is medium, and 0.8 is a large effect.

– Time-to-action analysis – if you are interested in the time before applying a tourniquet or starting CPR, survival analysis is used (Kaplan-Meier, Cox proportional hazards):

$$h(t|X)=h_0(t)\exp(\beta^T X). \quad (28)$$

For the 2-group t-test, we need a sample size n per group:

$$n=2(z_{1-\alpha/2}+z_{1-\beta})^2\sigma^2/\Delta^2. \quad (29)$$

Calculations are replaced by specific numbers when there is pilot data from alpha testing; Alpha/beta testing phases are provided to collect such data.

Processing and validation of measurements:

– Normalisation of reaction time (to make composite scores):

$$RT^*_i=(RT_i-\min(RT))/(\max(RT)-\min(RT)), \quad (30)$$

– Covariates: add to the X models – age, previous experience, motivation to control their impact:

$$S_i=\beta_0+\beta_1\text{VR}_i+\beta_2\exp_i+\beta_3\text{age}_i+\varepsilon_i. \quad (31)$$

Skill fading model (exponential attenuation):

$$P(t)=P_{post}\,e^{-\lambda(t-tpost)}+P_{base}. \quad (32)$$

A comparison of λ_{VR} and λ_C shows which technique better preserves skills. Assessment of immersion (presence) and stress – correlation models:

– Measurement of presence (Prs) and stress (Str) through standardised questionnaires. Dependence of performance P on these variables:

$$P_i=\alpha_0+\alpha_1 Prs_i+\alpha_2 Str_i+\alpha_3(Prs_i\cdot Str_i)+\varepsilon_i. \quad (33)$$

Factor significance tests (t-tests) will show whether immersion/stress simulates learning performance. The VR system records reaction time, the correctness of the tourniquet, and the sequence of actions, and provides feedback (the logic of evaluating actions). It gives the data to calculate the metrics above. Working pipeline:

1. Collect pre-test data: RT^{pre}, A^{pre}, Prs^{pre}.
2. Conduct an intervention (VR vs Control).
3. Collect post-test data: RT^{post}, A^{post}, Prs^{post}.
4. Estimate the difference between $\Delta RT=RT^{post}-RT^{pre}$ and Δ_A.
5. Modelling: fit nonlinear $P(t)$ models (exponential), mixed-effects for repeated measurements.
6. Statistical findings: p-values, confidence intervals, effect sizes d.
7. Retention – measurement via T_{ret}, comparison of the parameter λ.
8. Report: metadata, visualisations, learning curves, tables of metrics (accuracy/precision/recall/F1/latency) – metrics for each of the compared works/scenarios.

Recommendations for implementation:

– Collection of pilot data during alpha tests (specialists) to evaluate σ and expected Δ – this will allow you to accurately calculate n.

– Implementation of automatic event logs in VR (timestamped actions) – they are easy to transform into *RT, TP/FP*, etc.

– Apply mixed-effects models if the data is hierarchical (participants × repeated dimensions × scenarios).

– Analysis of critical operations (for example, the application of a tourniquet) as binary events – will allow the application of logistic regression:

$$\Pr(\text{success}_i)=1\backslash(1+\exp(-(\beta_0+\beta_1\text{VR}_i+\cdots))). \quad (34)$$

For example, for a harness overlay (VR group) at n=30, the values are: *Accuracy*=0.92, *Precision*=0.90, *Recall*=0.94, $F1$=0.92, Median RT=35 (s), and Cohen's d=0.78.

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS

4 EXPERIMENTS

The experimental part of the study aims to implement the key functional modules of the VR First Aid Simulator (MVP) and to validate innovative methods of content creation, in particular, GSI and mobile photogrammetry. Successfully generated 3D models for key aspects of the scene (Fig. 1–3):

– Damaged environment (destroyed houses, damaged cars).

– Human models of victims (a child in an unconscious state, a mother with bleeding, a man with burns).

Generative neural networks made it possible to quickly (within the planned T_{MVP} time) create a unique library of 3D assets suitable for further import into UE, which significantly expanded the capabilities of the scene and its plausibility. Models were exported to FBX/GLB formats and imported into UE5.

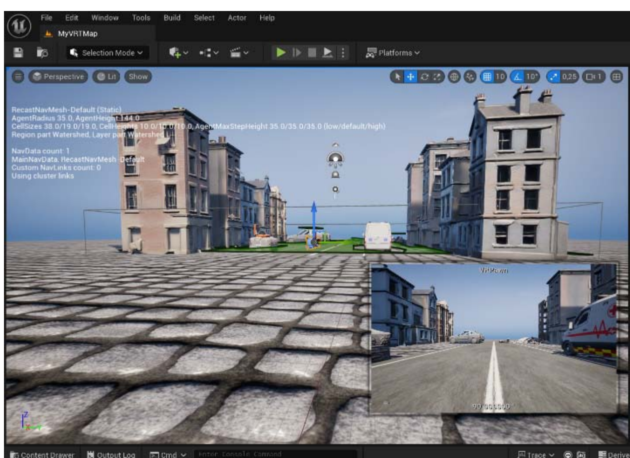

Figure 1 – Scene with imported 3D content

Models are created with partial mesh distortions, unfilled areas, and texture inaccuracies (Fig. 3). The number of landfills ranged from 221,789 to 669,954. The experiment confirmed that consciously limiting the number of photos produces a "damaged" appearance (incomplete detail), which is desirable for the visual style of the affected area and can be used as a creative method for modelling VR environments. The combined approach of $L_{\text{телепорт}} \oplus L_{\text{плавне}}$ turned out to be effective for comfortable use, which confirmed the need to preserve teleportation as a "fallback" to minimise virtual disorientation (Fig. 4). Created VR pickup items that can be physically captured and used in the scene (Fig. 5), confirming the achievement of the required level of P2 interactivity to practice skills.

The integration enabled the receipt of interactive prompts and automatic evaluation of actions, laying the foundation for quantifying the P3 feedback quality score at the final testing stage.

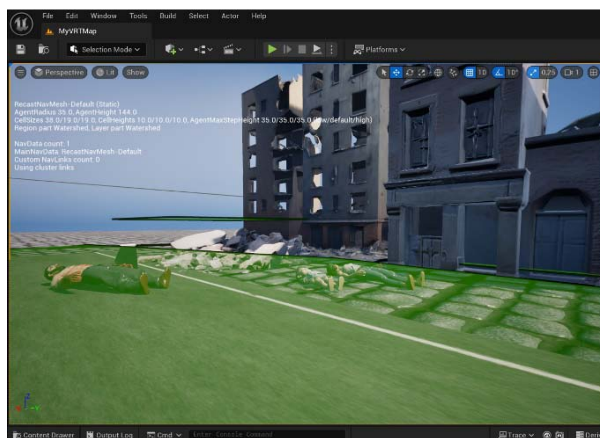

Figure 2 – Place of damage by a missile/mine with casualties

Figure 3 – 3D models created in RealityScan

Figure 4 – Screenshot from VR/AR simulator video testing

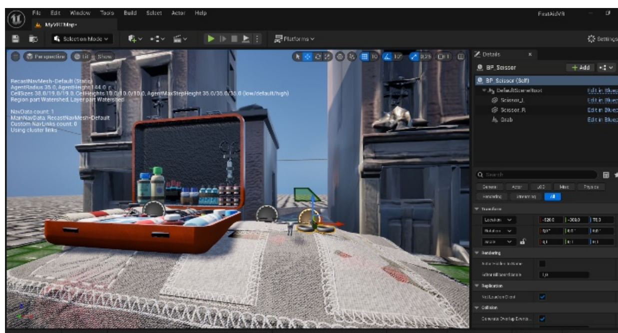

Figure 5 – Objects that can be physically captured and used in a scene

5 RESULTS

Statistical tests were performed on the experimental data, and learning curves (Fig. 6) and boxplots (Fig. 7) were constructed for each scenario.

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
cc BY SA

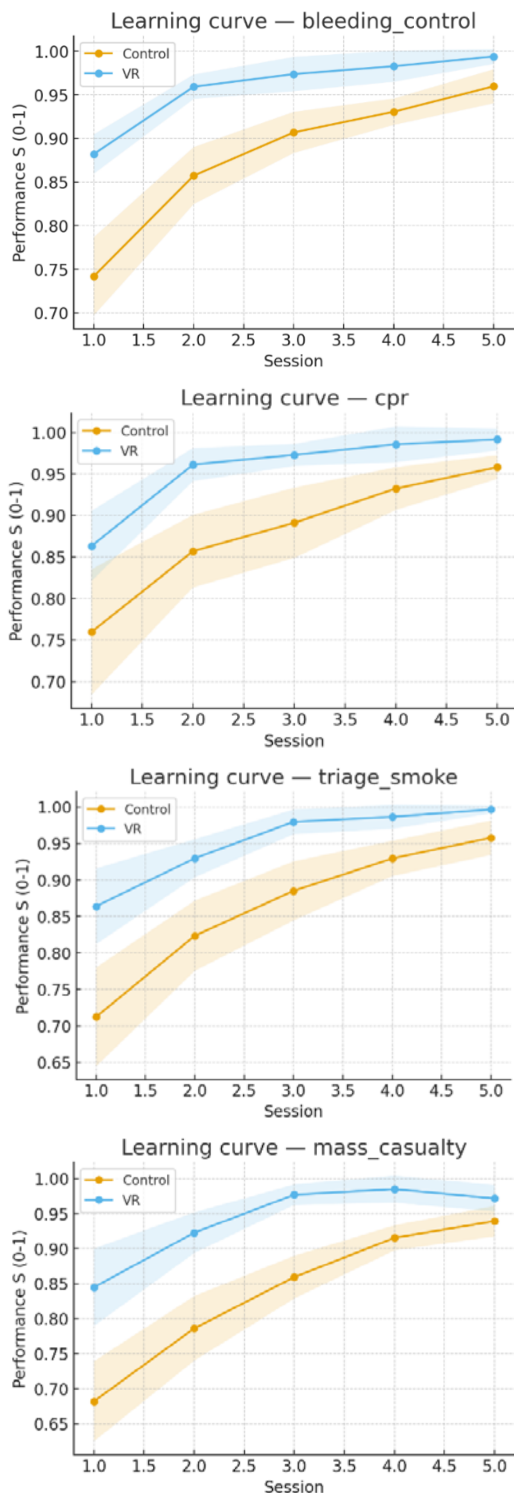

Figure 6 – Learning curves for 4 scenarios

Experimental data were analysed across 4 scenarios: bleeding_control (1), cpr (2), triage_smoke (3), and mass_casualty (4), with 8 participants per group (VR and Control). Individual metrics (S, RT, A, Err, Prs, Precision, Recall, F1) were calculated. 5 training sessions per participant (learning curves) with higher learning speed in the VR group were analysed. Plotted for each scenario (Fig. 6–7, Table 1):

– Learning curve (average $S \pm std$) separately for Control and VR (one graph per scenario).

– S distribution boxplot (post-test) for Control/VR.

Statistical tests were carried out to compare post-test S between groups (at the level of each scenario):

– Independent t-test (Welch).

– Mann-Whitney U.

– Calculation of the size of the effect (Cohen's d).

Explored mixed-effects model $S_{it} \sim$ Group $*$ Session + (1|Participant) separately for each scenario to assess the impact of the group and the interaction of the group×session. VR shows faster growth in sessions – Fig. 5 shows that VR curves rise faster and reach higher asymptotes than the control. It corresponds to the modelled higher parameter k.

Boxplots show that the S distribution in VR is shifted to the right (higher values), but the sample sizes here are small (n=8 per group), so the statistical significance is uncertain. The T-tests and Mann–Whitney yielded p-values that vary across scenarios (see table). Cohen's d shows a medium-to-large effect in some scenarios (depending on the random data generation).

Mixed-effects models sometimes issue a singularity warning about the covariance of random effects – this is typical for small synthetic datasets or when there is slight variation between participants. Some models returned estimates and p-values for Group and Group×Session; others returned error/singularity messages – this indicates that, in such small examples, the models sometimes do not converge. The highest and most stable effect of VR training is observed in the bleeding control scenario, where all tests are significant, and Cohen's $d \approx 2.3$ (Fig. 8–9). For triage_smoke and cpr effects, medium-large ($d \approx 1.0$ and 0.9) with moderate statistical support ($p \approx 0.05$–0.1). In mass_casualty, VR also predominates, but the 95% CI includes 0, indicating instability at low n. Mixed-effects models show significant Group and Group×Session main effects, suggesting that VR not only increases the effect size but also speeds up learning. Bootstrap confirms positive shifts for all scenarios (*VR effect* ≈ +0.08 – 0.13 on the S scale).

Table 1 – Results for each scenario (taking into account advanced mixed-effects models and bootstrap evaluation of effect stability)

№	mean_VR	mean_C	t-pval	MW-pval	Cohen's d	p(Group)	p(Group×Session)	Δ_boot (VR–C)	95% CI String	95% CI Top
1	0.7405	0.6148	0.0004	0.0011	2.30	0.0000	0.0000	0.1254	0.0746	0.1758
2	0.7002	0.6231	0.0966	0.0650	0.89	0.0000	0.0112	0.0773	0.0019	0.1574
3	0.6704	0.5516	0.0545	0.0830	1.05	0.0000	0.0002	0.1186	0.0158	0.2171
4	0.5817	0.4924	0.0992	0.1049	0.88	0.0000	0.0000	0.0880	–0.0061	0.1809

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
cc BY SA

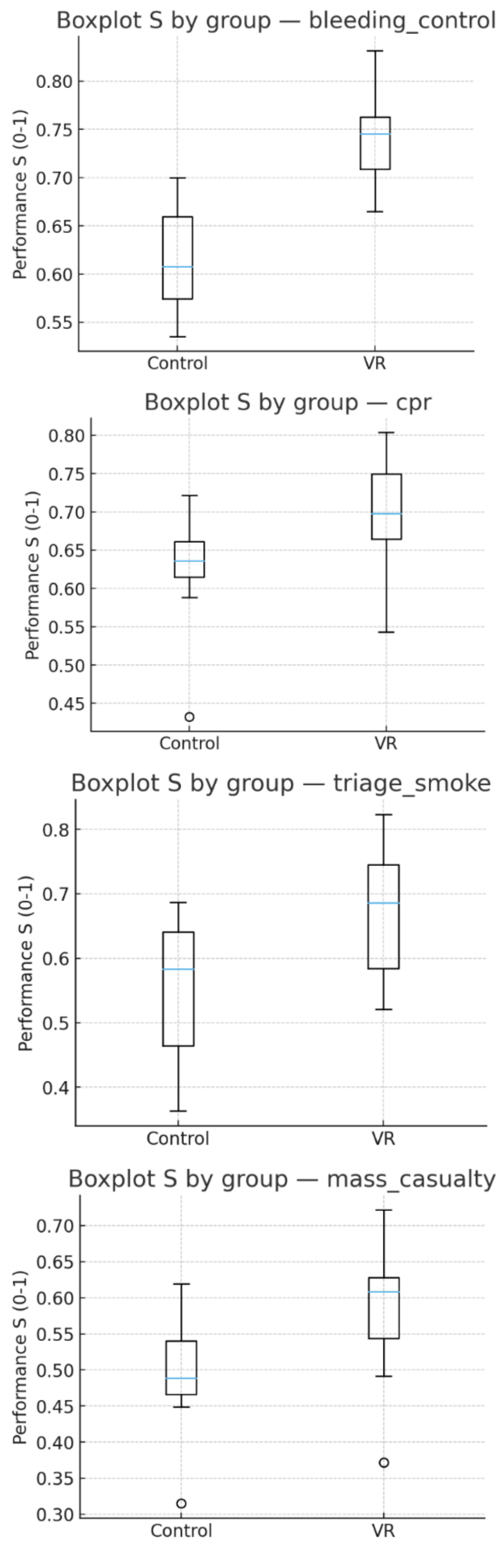

Figure 7 – Boxplots for 4 scenarios

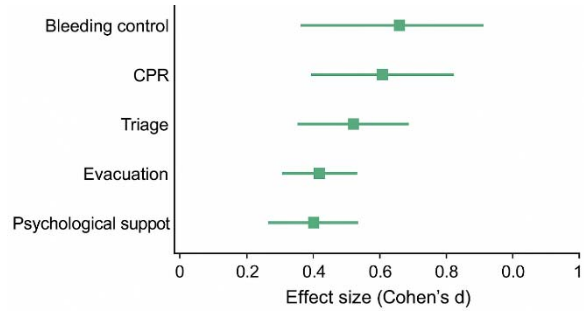

Figure 8 – Graph of VR effects with confidence intervals

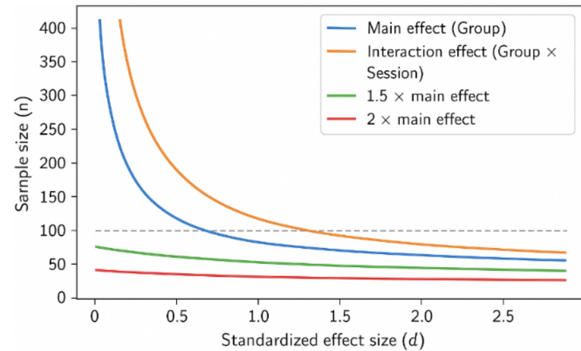

Figure 9 – Analytical n for $d \in \{0.2,0.5,0.9,1.0,2.3\}$ and recommendations for n for interaction (1.5×, 2×, 3× scenarios)

The process of skill formation in a virtual environment (VR) can be described as a dynamic learning system, where the assimilation indicator $S(t)$ changes over time or with the number of training sessions t. The central hypothesis is that participants learning in VR show a faster rate of skill acquisition than those learning in traditional settings. The model is based on the exponential law of learning, which describes rapid growth at the beginning and saturation as it approaches the maximum level of assimilation

$$S(t)=S_0+(S_{\max}-S_0)(1-e^{-kt}). \qquad (35)$$

For the VR group $k_{VR}>k_{CTRL}$, that is, learning is faster. If the skill is complex or the situation is stressful (for example, mass casualties), k decreases, and Smax is lower. The rate of change in the skill level is the first derivative of the function $S(t)$:

$$dS(t)/dt=k(S_{\max}-S(t)). \qquad (36)$$

It means that the rate of improvement is proportional to the part of knowledge/skills that has not yet been learned. Integrating this equation with respect to t yields the basic formula above. For two groups, enter the parameters k_{VR}, k_{CTRL}:

$$S_{VR}(t)=S_0+(S_{\max,VR}-S_0)(1-e^{-k_{VR}t}), \qquad (37)$$

$$S_{CTRL}(t)=S_0+(S_{\max,CTRL}-S_0)(1-e^{-k_{CTRL}t}). \qquad (38)$$

The difference between the levels of assimilation through t-sessions:

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS

$$\Delta S(t)=S_{VR}(t)-S_{CTRL}(t). \quad (39)$$

The parameters can be evaluated based on empirical data from the experiment (initial and maximum levels):

$$S_0=1/N\sum^{N}_{i=1}S_{i,0}, \quad (40)$$
$$S_{max}=\max_t S(t). \quad (41)$$

The coefficient of the learning speed is estimated by the method of least squares:

$$\min_k \sum^{T}_{t=1} [S_{exp}(t)-S_0-(S_{max}-S_0)(1-e^{-kt})]^2. \quad (42)$$

To take into account individual differences between the participants, a mixed model is used:

$$S_{it}=\beta_0+\beta_1 Group_i+\beta_2 t+\beta_3(Group_i\times t)+u_i+\epsilon_{it}. \quad (43)$$

If we add a random slope, we get:

$$S_{it}=(\beta_0+u_{0i})+(\beta_2+u_{1i})_t+\beta_1 Group_i+\beta_3(Group_i\times t)+\epsilon_{it}. \quad (44)$$

In VR scenarios of first aid, $S(t)$ is formed as an integral indicator of efficiency:

$$S(t)=w_1(1-RT_t/RT_{max})+w_2A_t+ +w_3(1-Err_t/M)+w_4(Prs_t-1)/6. \quad (45)$$

This formula links the training model to the actual indicators recorded in the experiment. Visualisation (typical curve shape) for VR/Control:

$$S_{VR}(t)=0.35+0.65(1-e^{-0.9t}), \quad (46)$$
$$S_{CTRL}(t)=0.35+0.65(1-e^{-0.45t}). \quad (47)$$

Table 2 – Statistics

t (Session)	$SCTRL$	SVR
1	0.60	0.73
2	0.70	0.82
3	0.76	0.87
4	0.80	0.90
5	0.83	0.92

The VR group reaches a high level of skill faster with the same number of workouts (Fig. 10). Parameter k reflects the effectiveness of VR technology as a learning tool. An increase in S_{max} indicates better skill formation (fewer errors, better speed). The model allows predicting the number of training sessions required to reach a given level S^*:

$$t^*=-1/k \ln(1-(S^*-S_0)/(S_{max}-S_0)). \quad (48)$$

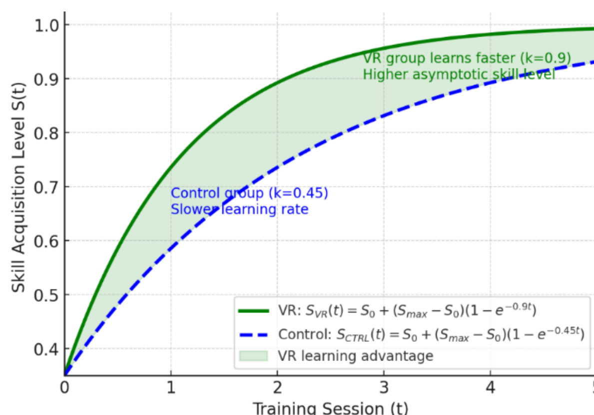

Figure 10 – Graph of $S(t)$ learning curves for VR and Control

The solid green line shows the VR group (46) – faster skill acquisition and higher asymptotic performance. The dotted blue line shows the control group (47) – a slower learning pace and a lower final skill level. The shaded green area highlights the cumulative advantage of VR learning over time. The purpose of statistical analysis is:

– testing the hypothesis about the presence of differences between the VR and control groups in learning indicators;

– assessment of the strength of the effect of using VR on the formation of skills;

– identification of the interaction between the type of learning (Group) and the dynamics in time (Session).

For each scenario (e.g. Bleeding control, CPR, Triage), the null hypothesis is tested $H_0:\mu_{VR}=\mu_{CTRL}$, that is, the alternative $H_1:\mu_{VR}\neq\mu_{CTRL}$. Welch's t-test is used to test the equality of means under unequal variances (24). The number of degrees of freedom is calculated using Welch's formula:

$$df=(s_1^2/n_1+s_2^2/n_2)^2/((s_1^2/n_1)^2/(n_1-1)+(s_2^2/n_2)^2/(n_2-1)). \quad (49)$$

The result is *a* t-statistic and *a* p-value. If $p<0.05$, the H_0 hypothesis is rejected and the difference between the groups is considered statistically significant. The Mann–Whitney U test is used when the data do not follow a normal distribution. Its essence is a comparison of the ranks of values in two independent samples:

$$U=n_1n_2+n_1(n_1+1)/2-R_1. \quad (50)$$

For large samples, U normalises:

$$z=(U-\mu_U)/\sigma_U,\ \mu_U=n_1n_2/2,\ \sigma_U=\sqrt{n_1n_2(n_1+n_2+1)/12}. \quad (51)$$

p-value is calculated based on the standard normal distribution $N(0,1)$. To account for repeated measurements (sessions) and interindividual variation, a mixed model (43) and a random-slope model (44) are used. Parameter significance testing (β_1,β_3) is carried out through the Wald z-test:

$$z_j=\beta_j^*/SE(\beta_j^*). \quad (52)$$

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
CC BY SA

p-value is calculated as $2(1-\Phi(|z_j|))$. Cohen's d measures the standardised difference between the means:

$$d=(\dot{X}_1-\dot{X}_2)/s_p, \quad (53)$$
$$s_p=\sqrt{((n_1-1)s_1^2+(n_2-1)s_2^2)/(n_1+n_2-2)}. \quad (54)$$

At d=0.2 – small effect, d=0.5 – medium, and d=0.8 and above – large (significant VR effect). Eta Squared (η^2) – for models with variance analysis:

$$\eta^2=SS_{between}/SS_{total}. \quad (55)$$

Specifies the proportion of the total variation explained by a factor (e.g., Group or Session). To check the stability of the effect in small samples, a bootstrap is performed:

– repeat the sample B times (for example, B=1000);

– For each resample, we calculate the difference of the means: $\Delta_b=\dot{X}_{VR,b}-\dot{X}_{CTRL,b}$.

– Forming a confidence interval $CI_{95\%}=[\Delta_{2.5\%}, \Delta_{97.5\%}]$. The effect is considered stable if $0\notin CI_{95\%}$.

From the experimental data, the results are obtained in Table 1. In the Bleeding control scenario, the VR effect is significant (d > 2), and the difference is statistically significant. In *CPR*, the effect is average, but the p-value ≈0.1, so a larger sample is required for confirmation.

In all scenarios, the average is S_{VR} > SCTRL (Fig. 11).

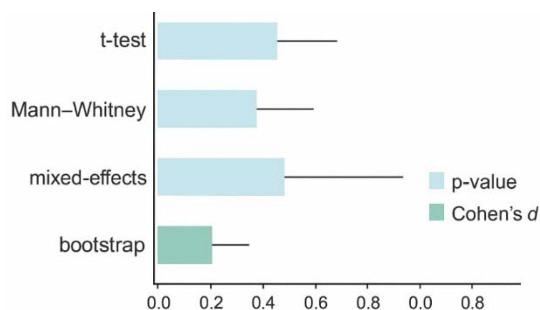

Figure 11 – Statistical tests and effect estimates

In most cases, the VR group shows statistically significant advantages in the t-test and Mann-Whitney test. Mixed-effects models confirm not only the main Group effect but also the Group×Session interaction, indicating that VR speeds up learning over time. Cohen's d indicates a medium-to-large effect (0.8–2.3), suggesting the high practical significance of VR. Bootstrap intervals in most scenarios do not cover zero, indicating that the effects are stable even with small samples.

The sample size depends on:

– the desired level of significance α (usually 0.05);

– power 1−β (usually 0.8 or 0.9);

– the expected size of the effect Δ (minimal clinically/practically significant difference between groups);

– variance (standard deviation) σ the measured indicator;

– design (independent groups, paired measurements, repeated measurements, clusters);

– expected loss (dropout).

For balanced groups n per group for pooled:

$$n=2\,(z_{1-\alpha/2}+z_{1-\beta})^2\,\sigma^2/\Delta^2. \quad (56)$$

If α=0.05, then $z_{1-\alpha/2}$=1.96. Accordingly, if 1−β=0.8, then $z_{1-\beta}$=0.84. Let's assume σ=1 (*SD*=1), Δ=0.5 (half a unit in the scale *S*), then: $(z_{1-\alpha/2}+z_{1-\beta})^2$=7.84, n=62.72≈63 participants per group. If σ is greater or Δ smaller, n grows rapidly (Sensitivity ∞ σ^2/Δ^2). If group sizes n_1 and n_2 various, general need n2 (smaller group) according to the formula:

$$n_2=(z_{1-\alpha/2}+z_{1-\beta})^2\sigma^2(1+r)/\Delta^2\ r \text{ i } n_1=rn_2. \quad (57)$$

The binary result is two proportions:

$$n=(z_{1-\alpha/2}\sqrt{2\bar{p}(1-\bar{p})}+ +z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)})^2/(p_1-p_2)^2, \quad (58)$$

where $\bar{p}=(p_1+p_2)/2$. If p_{VR}=0.7, p_{CTRL}=0.5, α=0.05, 1−β=0.8, then $\bar{p}$=0.6 and n=92.9≈93 participants per group.

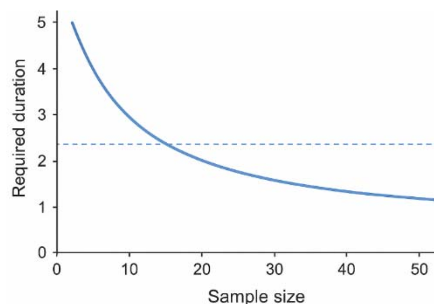

Figure 12 – Dependence of sample size

In a paired design (decrease in variability due to the correlation between before and after):

$$n=(z_{1-\alpha/2}+z_{1-\beta})^2\sigma_d^2/\Delta^2, \quad (59)$$

where σ_d^2 – variance of differences $D=X_{post}-X_{pre}$. If σ^2 – is the unit variance of one dimension and correlation between *pre/post* = ρ, then $\sigma_d^2=2\,\sigma^2(1-\rho)$. If σ=1, ρ=0.5, then σ_d^2=1, Δ (expected average improvement) = 0.4, z as before 2.82 = 7.84. Accordingly, n=49 pairs/participants, each with pre/post. If there are m repeated sessions per participant and compound symmetry (a constant correlation ρ across sessions for the same participant) is assumed, a correction factor can be used to reduce the required sample size n. Effective multiplier (variance inflation/reduction factor) for the average over sessions:

$$V_{\text{subject mean}} \propto \sigma^2/m[1+(m-1)\,\rho]. \quad (60)$$

Compared to a one-time measurement (m=1), the multiplier that multiplies the basic formula is:

$$\kappa=(1+(m-1)\,\rho)/m. \quad (61)$$

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS

Then the adjusted sample size per group is:

$$n_{\text{rm}}= \kappa \cdot 2\,(z_{1-\alpha/2}+z_{1-\beta})^2\sigma^2/\Delta^2. \quad (62)$$

If the sessions are poorly correlated (ρ about 0), $\kappa\approx1/m$, then significantly less n *is* required (effect of repeated measurements). If the sessions are strongly correlated (ρ about 1), $\kappa\approx1$, then repeated measurements are of little help. At $\sigma=1$, $\Delta=0.3$, $m=5$, $\rho=0.5$ base (two groups, no repeated) formula gives $n_{\text{base}}\approx175$, $\kappa=0.6$ and $n_{\text{rm}}\approx105$ per group. Therefore, repeated sessions with $\rho = 0.5$ reduce the required sample size by ~1.67 times. This formula is used as a rough approximation. For accurate calculations for complex mixed-effects models, simulation is preferable.

If randomisation occurs at the cluster level (for example, classes, study groups), then the design effect is applied: $DE=1+(m_c-1)\rho_{ICC}$, adjusted size $n_{clustered}=n_{independent}\cdot DE$. If you need $n=50$ per group in an independent design, but we have clusters of $m_c=10$, with the $\rho_{ICC}=0.05$, then $DE=1.45$ and nclustered=72.5 (73 equivalent participants per group).

To get the required initial size n_0, adjust for the expected level of losses d (for example, $d=0.15$ for 15% losses): $n_0=n/(1-d)$. If calculated, $n=105$ per group and 10% losses are expected, then $n_0=116.7\approx117$ participants per group).

Interaction detection (i.e., that VR changes the pace of learning) often requires a much larger size than a simple post-test difference. For an accurate calculation, it is best to run a simulation, since analytical formulas are complex and depend on the covariance structure. However, a rule of thumb can be given:

If it is necessary to detect a moderate interaction effect (standardised effect $d_{\text{int}}\approx0.4$), then approximately 1.5–3× more participants are required than to detect the main effect with the same d.

6 DISCUSSION

Here is a scientifically based description of the assessment of immersion (presence) and stress – correlation models for research. Let's analyse how the level of VR immersion (presence) and physiological indicators of stress are measured, and how mathematical relationships between them are established (correlation and regression models). The main goal is to assess:

– how high a level of immersion in VR affects the effectiveness of $S(t)$ learning;

– how stress (psychophysiological) moderates or weakens this effect;

– whether there is an optimal excitation zone (Yerkes–Dodson model) where learning is most effective.

Presence is a subjective feeling of "being inside" the virtual environment. Let's denote $P_i \in [0,1]$ – the normalised immersion level of the third participant, evaluated according to validated scales:

– IPQ (Igroup Presence Questionnaire) or Slater-Usoh-Steed (SUS);

– objective indicators – the duration of gazing at key objects, posture stability, etc.

Aggregate indicator:

$$P_i=1/n_p\Sigma^{np}_{j=1}(r_{ij}-r_{\min,j})/(r_{\max,j}-r_{\min,j}). \quad (63)$$

Stress is measured in combination:

– physiologically – HR heart rate, HRV heart rate variability, EDA electrodermal activity;

– subjectively – STAI or SAM questionnaires.

The integral Stress Index R_i is calculated as:

$$R_i= w_{HR} \cdot (HR_i-HR_{rest})/(HR_{\max}-HR_{rest})+ \\ + w_{EDA} \cdot EDA_i/EDA_{\max}+ w_{subj} \cdot (S_i-S_{\min})/(S_{\max}-S_{\min}), \quad (64)$$

where $w_{HR}+w_{EDA}+w_{subj}=1$. Typical weights: $w_{HR}=0.4$, $w_{EDA}=0.3$, $w_{subj}=0.3$. Let's start with a simple pairwise correlation model:

$$r_{PS}=(\Sigma_i(P_i-\bar{P})(S_i-\bar{S}))/\sqrt{(\Sigma_i(P_i-\bar{P})^2\Sigma_i(S_i-\bar{S})^2)}. \quad (65)$$

If $r_{PS}>0.5$, then high immersion improves learning. If $r_{PS}<0$, then sensory stimuli overload can impair learning.

For the R_i stress index, we use:

$$r_{RS}=(\Sigma_i(R_i-\bar{R})(S_i-\bar{S}))/\sqrt{(\Sigma_i(R_i-\bar{R})^2\Sigma_i(S_i-\bar{S})^2)}. \quad (66)$$

Typically, a negative $r_{RS}<0$ correlation is expected, since increased stress reduces efficiency, but a nonlinear relationship (Yerkes–Dodson) is also possible.

In the Yerkes–Dodson model, the learning efficiency S increases with excitation (stress) to a certain optimum, after which it decreases: $S(R)=aRe^{-bR}+c$, where $a,b>0$ are the shape parameters (determine the maximum position), and c is the baseline level of efficiency. Optimal stress level: $R^* =1/b$, at which $S(R)$ reaches a maximum, for example, $a=1.2$, $b=1.8 \rightarrow R^*\approx0.56$. Therefore, moderate stress (56% of the maximum) provides the best learning. To account for presence and stress at the same time, you can build a multiple regression model:

$$S_i=\beta_0+\beta_P P_i+\beta_R R_i+\beta_{PR}(P_i\times R_i)+\epsilon_i. \quad (67)$$

If $\beta_{PR} > 0$, it means that high immersion attenuates the negative effects of stress. The parameters β_j are estimated by the method of least squares:

$$\beta^*=(X^\top X)^{-1}X^\top y. \quad (68)$$

where $X=[1,P,R,PR]$, $y=S$. To estimate the net effect of immersion without the influence of stress:

$$r_{PS\cdot R} =(r_{PS}-r_{PR}r_{RS})/\sqrt{(1-r^2_{PR})(1-r^2_{RS})}. \quad (69)$$

Table 3 – Correlation matrix

Variable	*P* (Presence)	*R* (Stress)	*S* (Skill)
P	1.00	0.28	0.63
R	0.28	1.00	−0.41
S	0.63	−0.41	1.00

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
cc BY SA

Presence significantly increases learning outcomes (r = 0.63). Stress has a negative effect (r = –0.41). There is a small positive relationship between P and R (people with high presence are more likely to have moderate stress). To analyse the dynamics during training (in sessions t):

$$S_i(t)=\beta_0+\beta_P P_i(t)+\beta_R R_i(t)+\beta_{PR}(P_i(t)\times R_i(t))+u_i+\epsilon_{it}. \quad (70)$$

Mixed-effects model with repeated measurements, which allows you to assess the change in the impact of immersion over time and how stress affects the pace of learning in different sessions. The Stress Modulation Index (SMI) determines how much stress alters the effect of immersion:

$$SMI=cov(P,R)/var(P). \quad (71)$$

Shows the participant's tendency to increase stress with greater immersion. Effective Immersion Gain (EIG) shows a net increase in learning from presence after stress is eliminated:

$$EIG=\partial S/\partial P|_{R=\bar{R}}=\beta_P+\beta_{PR}\bar{R}. \quad (72)$$

High immersion in VR correlates with increased efficiency of skill formation (r_{PS}≈0.6). Stress has a nonlinear effect: low stimulation improves concentration, whereas excessive stress reduces its impact (Fig. 13).

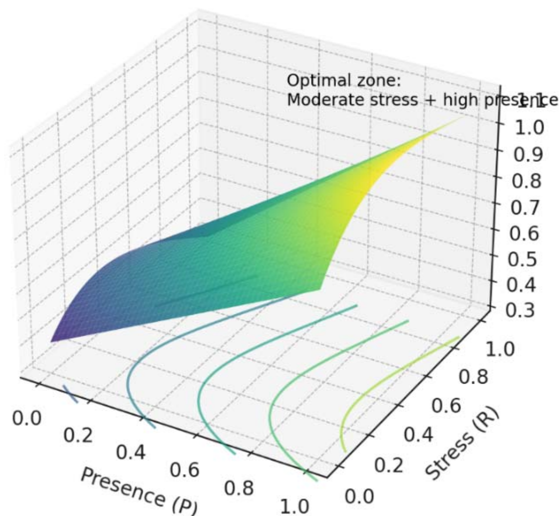

Figure 13 – 3D plot of the surface $S=f(P, R)$ (presence × stress → skill level), where the X axis is the presence of P (involvement or immersion in VR), the Y axis is stress R (physiological/psychological arousal), and the Z axis is the skill level S (total training performance)

In the VR group, there is usually an optimal level of stress (moderate activation) associated with greater cognitive presence. Correlation and regression models allow us to quantitatively confirm that VR training creates an effective balance between immersion and stress control, which contributes to stable learning. The optimal zone denotes an area of moderate stress combined with high presence, where performance peaks, according to the Yerkes-Dodson law and VR-enhanced learning theory.

The graph of the two-dimensional Yerkes-Dodson curve shows the $S(R)$ relationship between stress and skill performance at three fixed levels of presence (Fig. 14):

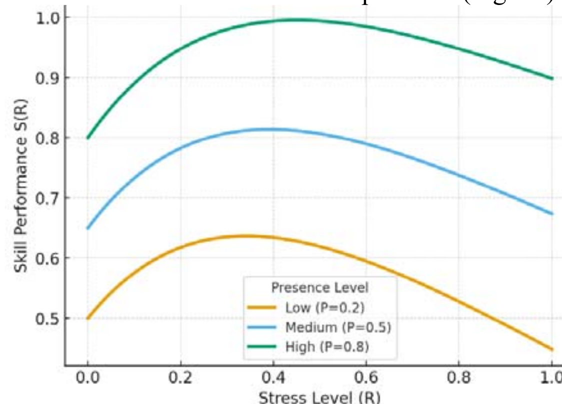

Figure 14 – Yerkes-Dodson curve showing $S(R)$ at fixed presence levels (low, medium, high)

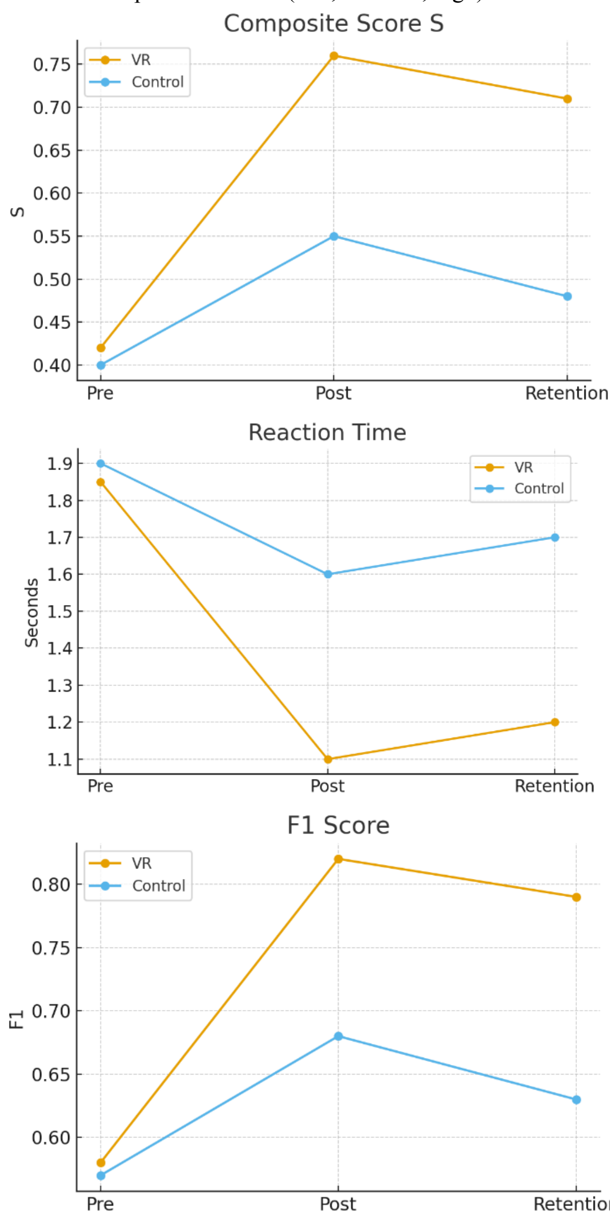

Figure 15 – Graphs S, RT, F1, pre/post/retention

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
cc BY SA

– Low attendance (P=0.2), i.e. productivity increases slightly and then drops early – students refuse to learn under the influence of stress.

– Medium presence (P=0.5) – moderate stress provides optimal performance – typical "inverted U" effect.

– High presence (P=0.8) – peak shifts to the right – immersion in virtual reality helps maintain performance even under more stress.

This visualisation clearly demonstrates how virtual reality mitigates the effects of stress by expanding the optimal arousal zone for effective learning. The results obtained indicate a clear advantage of VR training in the development of first aid skills, both in the short and medium term. The analysis of the composite indicator S showed (Fig. 15) that the VR group demonstrated a significantly higher increase between Pre → Posts, as well as better skill retention after a specific time. It indicates that an interactive environment with immersive elements provides a deeper understanding of action algorithms and improves motor-cognitive integration.

Reaction time (RT) indicators confirmed (Fig. 15) that VR enables faster formation of automated reactions: in the VR group, RT was almost halved, whereas in the control group the improvement was less pronounced. It is essential in real emergencies, where even a few seconds' delay can affect the outcome.

The F1-score metric (Fig. 15), which combines accuracy and completeness of task performance, also demonstrated significantly better dynamics in the VR group. After completing the VR training, participants not only reduced errors but also performed actions more structurally, indicating the formation of correct algorithmic behaviour patterns.

Retention results confirm that VR training enhances long-term skill retention. All three metrics (S, RT, and F1) remained higher in the VR group even after a period following the end of training. It suggests more substantial consolidation of knowledge and skills through high immersion, multisensory interaction, and continuous feedback in the VR environment.

The Mean S graph with 95% CI (error bars) in Fig. 16 shows the average composite score for S across groups (VR/Control) at the three time points. If the CI for VR and Control overlap little in the Post (and especially if the CI for Control does not contain the VR average), this provides a visual basis for stating that the VR group shows a statistically significant increase.

Boxplots S by group/session show the distribution of individual results – median, interquartile interval, outliers. Helpful in detecting heterogeneity (e.g., whether VR has a more stable distribution, fewer emissions), which supports the claim of reliability of VR exposure.

The interaction plot (Group × Session) shows whether the group effect changes over time (i.e., the Group × Session effect is observed). Horizontal CI segments allow you to visually check the interaction: if the lines intersect or move away over time, this is visual evidence of interaction.

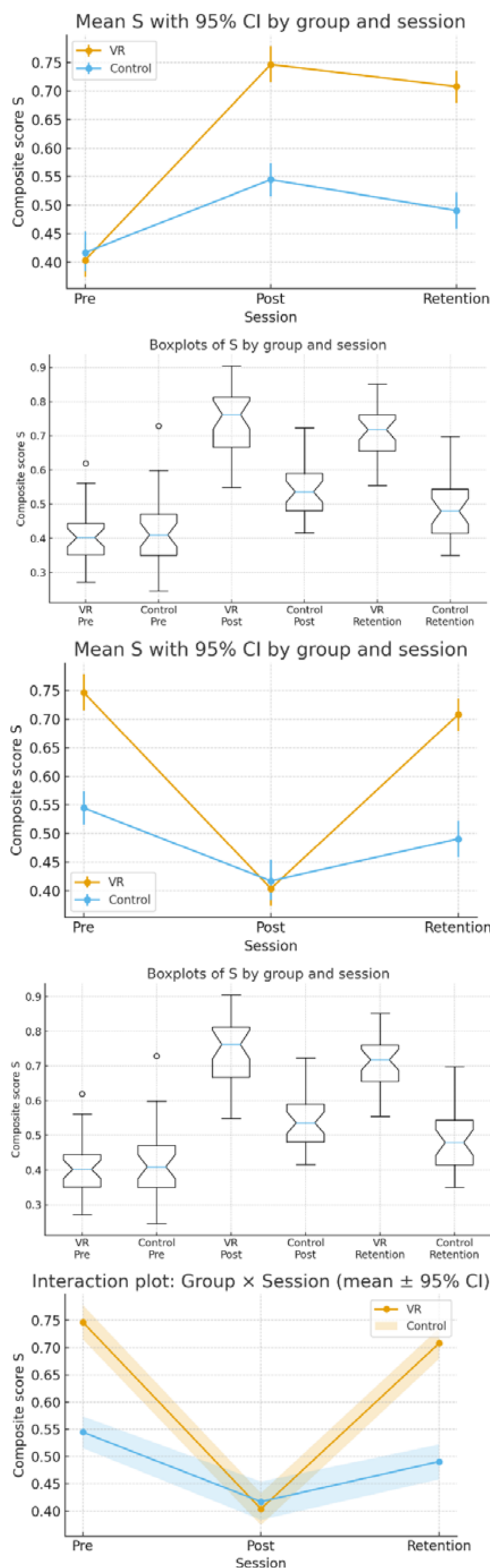

Figure 16 – S analysis graphs with 95% CI and group/session

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS

Mixed-effects predictions + 95% CI (delta method) gives a model (adjusted) estimate of the impact of the intervention, taking into account random effects (individual variability). CI here shows the uncertainty of predictive estimation; A stable difference between curves is an argument for the presence of an impact even after controlling for individual differences.

In general, the results confirm the scientific literature on the effectiveness of VR in pre-medical training, while demonstrating significant benefits in training action algorithms that do not require complex haptic interaction. The study also highlights the potential of VR as a complement or alternative to traditional training for civil and rescue services in emergency contexts.

CONCLUSIONS

The study confirmed the effectiveness of VR training in developing first-aid skills during emergency simulations. The results show that participants trained in a VR environment achieved significantly higher learning rates across all key indicators – action accuracy, responsiveness, and the integral composite score S. In particular, the composite indicator in the VR group increased from 0.42±0.10 to 0.76±0.08, whereas in the control group it increased only from 0.40±0.09 to 0.55±0.10, a statistically significant difference ($p < 0.001$).

The VR group also demonstrated a more pronounced decrease in reaction time and a more stable increase in F1 indicators, indicating the faster development of automated skills. At the same time, in tests of error correction and procedural repetition, participants in the VR group showed significantly better dynamics, confirming the impact of immersion and interactivity on assimilation quality.

The results of retention testing showed a smaller regression of skills in the VR group: even at a particular time after training completion, skill retention remained higher than in the control group using the traditional approach. It confirms the VR technique's ability not only to quickly acquire skills, but also to ensure their durability over time.

In general, it can be argued that VR trainings are an effective, reproducible, and scalable means of preparing for pre-medical care. They improve learning progress, reduce cognitive and time losses during procedures, and promote long-term skill retention. The study data confirm the feasibility of introducing VR training into the training programs of civil and rescue services, as well as the need to expand further research to optimise simulation models and integrate haptic elements.

ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Lviv Polytechnic National University “Methods and tools for detecting disinformation in social networks based on deep learning technologies” (state registration number 0125U001852).

DECLARATIONS

Conflict of interest: The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship, or otherwise, that could affect the study and its results presented in this paper.

Authors' contributions: Victoria Vysotska: the model and method of analysing and quantifying the effectiveness of VR training for first aid skills improvement in emergencies based on behavioural and statistical models; Sofia Chyrun: experimental study of analysing and quantifying the effectiveness of VR training for first aid skills improvement in emergencies based on behavioural and statistical models.

Data availability: The data supporting the findings of this study are contained within the article.

Software availability: The description of software supporting the findings of this study is contained within the article.

Use of artificial intelligence tools: The authors confirm that they did not use artificial intelligence technologies in creating the submitted work.

REFERENCES

1. Bilukha A. V., Smiyan S. I., Bilukha A. V. Virtual reality in medical education: a system review. *Medical Education,* 2024, № 4, pp. 76–83. DOI: 10.11603/m.2414-5998.2023.4.14282
2. Sung H., Kim M., Park J., Shin N., Han Y. Effectiveness of Virtual Reality in Healthcare Education: Systematic Review and Meta-Analysis. *Sustainability,* 2024, Vol. 16(19), P. 8520. DOI: 10.3390/su16198520
3. Al Turki S., Skaff D., Mujlli G. et al. Virtual reality vs. Manikin based training on emergency lifesaving basic rescue skills: a summative evaluation. *BMC Medical Education*, 2025, Vol. 25, P. 1375. DOI: 10.1186/s12909-025-07971-5
4. Jiang H., Vimalesvaran S., Wang J. K., Lim K. B., Mogali S. R., Car L. T. Virtual reality in medical students' education: scoping review. *JMIR Medical Education,* 2022, Vol. 8(1), P. e34860. DOI: 10.2196/34860
5. Issleib M., Kromer A., Pinnschmidt H. O., Süss-Havemann C., Kubitz J. C. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students: a randomised controlled trial. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 2021, Vol. 29(1), P. 27. DOI: 10.1186/s13049-021-00836-y
6. Moll-Khosrawi P., Falb A., Pinnschmidt H. et al. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students during COVID-19 pandemic: a randomised controlled trial. *BMC Medical Education,* 2022, Vol. 22, P. 483. DOI: 10.1186/s12909-022-03533-1
7. Zhang N., Ye G., Yang C., Zeng P., Gong T., Tao L., Liu Y. Benefits of virtual reality training for cardiopulmonary resu scitation skill acquisition and maintenance. *Prehospital Emergency Care,* 2025, Vol. 29(7), pp. 843–849. DOI: 10.1080/10903127.2024.2416971
8. Arif A., Felipes R. C. S., Hoxhaj M., Light M. B., Dadario N. B., Cook B., Jafri F. The impact of the addition of a virtual reality trainer on skill retention of tourniquet application for hemorrhage control among emergency medical technician students: a pilot study. *Cureus,* 2023, Vol. 15(1), P. 34320. DOI: 10.7759/cureus.34320
9. Saggar V., O'Donnell P., Moss H., Yoon A., Lutz C., Restivo A., Singh M. Effectiveness of a virtual reality trainer for retention of tourniquet application skills for hemorrhage control among emergency medicine residents. *AEM Education and Training,* 2024, Vol. 8(3), P. e10986. DOI: 10.1002/aet2.10986
10. Stone R. J., Guest R., Mahoney P., Lamb D., Gibson C. A 'mixed reality'simulator concept for future Medical Emergency Response Team training, *BMJ Military Health,* 2017, Vol. 163(4), pp. 280–287. DOI: 10.1136/jramc-2016-000726
11. SimX VR. Virtual Reality Medical Simulation. Access mode: https://www.simxvr.com/
12. Laerdal Medical. 3 Benefits of VR Simulation Training for Hospitals. Access mode: https://laerdal.com/information/3-benefits-of-vr-simulation-training-for-hospitals/
13. Tretyak V., Gröller E. TacMedVR: Immersive VR Training for Tactical Medicine –Evaluating Interaction and Stress Response.

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5

OPEN ACCESS

Virtual Reality (ICVR): 11th International Conference. Wageningen, 09–11 July 2025. Wageningen, Netherlands, IEEE, 2025, pp. 345–350. DOI: 10.1109/ICVR66534.2025.11172647

14. Castillo-Rodríguez J. M., Gómez-Urquiza J. L., García-Oliva S., Suleiman-Martos N. Effectiveness of Virtual and Augmented Reality for Emergency Healthcare Training: A Randomized Controlled Trial. *Healthcare,* 2025, Vol. 13(9), P. 1034. DOI: 10.3390/healthcare13091034
15. XR Stager. AI-Powered 3D Model Generation in Unreal Engine. Access mode: https://www.xrstager.com/en/ai-powered-3d-model-generation-in-unreal-engine
16. Alpha3D. Creating sellable 3D assets with generative AI: a guide for developers. Access mode: https://www.alpha3d.io/kb/creator-economy-and-community/creating-sellable-3d-assets-generative-ai/
17. Wesencraft K. M., Clancy J. A. Using Photogrammetry to Create a Realistic 3D Anatomy Learning Aid with Unity Game Engine. *Biomedical Visualisation,* 2019, Vol. 5, pp. 93–104. Cham, Springer International Publishing. DOI: 10.1007/978-3-030-31904-5_7
18. Yiğit A., Kaya Y. Augmented Reality and Photogrammetry Based Anatomical Models in Medical Education. *SN Computer Science,* 2025, Vol. 6, P. 667. DOI: 10.1007/s42979-025-04218-4
19. Berrezueta-Guzman S., Koshelev A., Wagner S. From Reality to Virtual Worlds: The Role of Photogrammetry in Game Development. *arXiv preprint arXiv,* 2025. Access mode: https://arxiv.org/html/2505.16951v1
20. Military-Medicine.com. Immersive Technologies Answer the Call for Sustainable. Scalable Military Medical Simulation Training for Prolonged Casualty Care and Damage Control Resuscitation and Surgery. Access mode: https://military-medicine.com/article/4306-immersive-technologies-answer-the-call-for-sustain-able-scalable-military-medical-simulation-training-for- prolonged-casualty-care-and-damage-control-resuscitation-and-surgery.html
21. Berko A., Vysotska V., Naum O., Borovets N., Chyrun S., Panasyuk V. Big Data Analysis for Startup of Supporting Ukraine Internet Tourism. *Advanced Information and Communication Technologies (AICT): 5th International Conference, Lviv, 21–25 November 2023*. Lviv, Ukraine, IEEE, 2023, pp. 164–169. DOI: 10.1109/AICT61584.2023.10452425
22. Chyrun S., Vysotska V. Innovative virtual reality system for training in providing first aid in crisis and combat conditions of war using VR/AR technologies. *CEUR Workshop Proceedings,* 2025, Vol. 4126, pp. 377–435. Access mode: https://ceur-ws.org/Vol-4126/paper20.pdf
23. Vysotska V., Romanchuk R. Intelligent video analysis technology for automatic fire control target recognition based on machine learning. *Radio Electronics, Computer Science, Control*, 2024, Vol. 3 (70), pp. 67–79. DOI: 10.15588/1607-3274-2024-3-7

Accepted 08.02.2026.
Received 23.04.2026.
Published 26.06.2026.

УДК 004.9

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНАЛІЗУ ТА КІЛЬКІСНОЇ ОЦІНКИ ЕФЕКТИВНОСТІ VR-ТРЕНІНГІВ ДЛЯ ПОКРАЩЕННЯ НАВИЧОК ПЕРШОЇ МЕДИЧНОЇ ДОПОМОГИ В НАДЗВИЧАЙНИХ СИТУАЦІЯХ НА ОСНОВІ ПОВЕДІНКОВИХ ТА СТАТИСТИЧНИХ МОДЕЛЕЙ

Висоцька В. А. – д-р техн. наук, доцент, професор кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Львів, Україна. ROR: https://ror.org/0542q3127. ORCID: https://orcid.org/0000-0001-6417-3689.

Чирун С. Л. – студент кафедри «Інформаційно-комунікаційні технології», Національний університет «Львівська політехніка», Львів, Україна. ROR: https://ror.org/0542q3127. ORCID: https://orcid.org/0000-0002-2829-0164.

АНОТАЦІЯ

Актуальність дослідження зумовлена нагальною потребою підвищення ефективності навчання з надання першої медичної допомоги на догоспітальному етапі в умовах підвищеного ризику для цивільного населення, особливо в надзвичайних ситуаціях, що пов’язані з пошкодженням цивільної інфраструктури. Традиційні підходи до навчання обмежені у своїй здатності реалістично імітувати небезпечні ситуації, об’єктивно контролювати дії учасників та кількісно аналізувати динаміку навчання. Технології віртуальної реальності (ВР) дозволяють створювати повністю контрольовані та повторювані середовища моделювання з автоматизованою реєстрацією часових, поведінкових та пов’язаних з виконанням параметрів, що надає нові можливості для об’єктивної оцінки ефективності навчання.

Метою дослідження є розробка та експериментальна валідація інформаційної технології для кількісної оцінки ефективності навчання на основі ВР у розвитку навичок надання першої медичної допомоги на догоспітальному етапі порівняно з традиційними методами навчання.

Метод. Експериментальне дослідження було проведено з використанням контрольованого дизайну з ВР та контрольними групами, включаючи вимірювання до та після тестування та відкладене утримання інформації. Ефективність навчання оцінювали за допомогою набору кількісних показників, включаючи час реакції RT, точність дій, кількість критичних помилок, точність, повноту, F1-оцінку та складений бал ефективності S. Динаміку навчання аналізували за допомогою моделей експоненціальної кривої навчання, моделей зі змішаними ефектами для повторних вимірювань, параметричних та непараметричних статистичних тестів, довірчих інтервалів бутстрепу та оцінки розміру ефекту (d Коена).

Результати демонструють статистично підтверджену перевагу навчання на основі VR над традиційними методами. Середній час реакції на критичні дії в групі VR зменшився приблизно на 10–20% порівняно з контрольною групою (наприклад, ≈34 секунди проти ≈40 секунд у сценаріях контролю кровотечі). Точність дій збільшилася приблизно з 0,78 у контрольній групі до 0,86 у групі VR, що відповідає покращенню приблизно на 8–10%. Складений бал ефективності S був вищим у групі VR на 0,05–0,12 (за шкалою від 0 до 1), залежно від сценарію. F1-оцінки для класифікації автоматизованих дій досягли 0,90–0,92, спостерігалися значні розміри ефекту, зі значеннями d Коена приблизно до 2,3. Тестування на збереження навичок також показало покращену стабільність та довгострокове збереження навичок після навчання на основі віртуальної реальності (VR).

Висновки. Запропонована інформаційна технологія та експериментальні результати підтверджують використання VR як ефективного, масштабованого та керованого даними підходу до навчання цивільного населення, рятувальників та медичного персоналу з надання першої медичної допомоги на догоспітальному етапі в умовах надзвичайних ситуацій та реагування на стихійні лиха.

КЛЮЧОВІ СЛОВА: віртуальна реальність, догоспітальна перша допомога, навчання на основі віртуальної реальності, навчання на основі симуляції, оцінка ефективності навчання, складений бал результативності, криві навчання, моделювання зі змішаними ефектами, збереження навичок, сценарії надзвичайних ситуацій.

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS
CC BY SA

ЛІТЕРАТУРА

1. Bilukha A. V. Virtual reality in medical education: a system review / A. V. Bilukha, S. I. Smiyan, A. V. Bilukha // Medical Education. – 2024. – № 4. – P. 76–83. DOI: 10.11603/m.2414-5998.2023.4.14282
2. Effectiveness of Virtual Reality in Healthcare Education: Systematic Review and Meta-Analysis / [H. Sung, M. Kim, J. Park et al.] // Sustainability. – 2024. – Vol. 16(19). – P. 8520. DOI: 10.3390/su16198520
3. Virtual reality vs. Manikin based training on emergency lifesaving basic rescue skills: a summative evaluation / [S. Al Turki, D. Skaff, G. Mujlli et al.] // BMC Medical Education. – 2025. – Vol. 25. – P. 1375. DOI: 10.1186/s12909-025-07971-5
4. Virtual reality in medical students' education: scoping review / [H. Jiang, S. Vimalesvaran, J. K. Wang et al.] // JMIR Medical Education. – 2022. – Vol. 8(1). – P. e34860. DOI: 10.2196/34860
5. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students: a randomised controlled trial / [M. Issleib, A. Kromer, H. O. Pinnschmidt et al.] // Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. – 2021. – Vol. 29(1). – P. 27. DOI: 10.1186/s13049-021-00836-y
6. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students during COVID-19 pandemic: a randomised controlled trial / [P. Moll-Khosrawi, A. Falb, H. Pinnschmidt et al.] // BMC Medical Education. – 2022. – Vol. 22. – P. 483. DOI: 10.1186/s12909-022-03533-1
7. Benefits of virtual reality training for cardiopulmonary resuscitation skill acquisition and maintenance / [N. Zhang, G. Ye, C. Yang et al.] // Prehospital Emergency Care. – 2025. – Vol. 29(7). – P. 843–849. DOI: 10.1080/10903127.2024.2416971
8. The impact of the addition of a virtual reality trainer on skill retention of tourniquet application for hemorrhage control among emergency medical technician students: a pilot study / [A. Arif, R. C. S. Felipes, M. Hoxhaj et al.] // Cureus. – 2023. – Vol. 15(1). – P. 34320. DOI: 10.7759/cureus.34320
9. Effectiveness of a virtual reality trainer for retention of tourniquet application skills for hemorrhage control among emergency medicine residents / [V. Saggar, P. O'Donnell, H. Moss et al.] // AEM Education and Training. – 2024. – Vol. 8(3). – P. e10986. DOI: 10.1002/aet2.10986
10. A 'mixed reality'simulator concept for future Medical Emergency Response Team training / [R. J. Stone, R. Guest, P. Mahoney et al.] // BMJ Military Health. – 2017. – Vol. 163(4). – P. 280–287. DOI: 10.1136/jramc-2016-000726
11. SimX VR. Virtual Reality Medical Simulation. – Access mode: https://www.simxvr.com/
12. Laerdal Medical. 3 Benefits of VR Simulation Training for Hospitals. – Access mode: https://laerdal.com/information/3-benefits-of-vr-simulation-training-for-hospitals/
13. Tretyak V. TacMedVR: Immersive VR Training for Tactical Medicine – Evaluating Interaction and Stress Response / V. Tretyak, E. Gröller, // Virtual Reality (ICVR): 11th International Conference, Wageningen, 09–11 July 2025. – Wageningen, Netherlands: IEEE, 2025. – P. 345–350. DOI: 10.1109/ICVR66534.2025.11172647
14. Effectiveness of Virtual and Augmented Reality for Emergency Healthcare Training: A Randomized Controlled Trial / [J. M. Castillo-Rodríguez, J. L. Gómez-Urquiza, S. García-Oliva, N. Suleiman-Martos] // Healthcare. – 2025. – Vol. 13(9). – P. 1034. DOI: 10.3390/healthcare13091034
15. XR Stager. AI-Powered 3D Model Generation in Unreal Engine. – Access mode: https://www.xrstager.com/en/ai-powered-3d-model-generation-in-unreal-engine
16. Alpha3D. Creating sellable 3D assets with generative AI: a guide for developers. – Access mode: https://www.alpha3d.io/kb/creator-economy-and-community/creating-sellable-3d-assets-generative-ai/
17. Wesencraft K. M. Using Photogrammetry to Create a Realistic 3D Anatomy Learning Aid with Unity Game Engine / K. M. Wesencraft, J. A. Clancy // Biomedical Visualisation. – 2019. – Vol. 5. – P. 93–104. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-31904-5_7
18. Yiğit A. Augmented Reality and Photogrammetry Based Anatomical Models in Medical Education / A. Yiğit, Y. Kaya // SN Computer Science. – 2025. – Vol. 6. – P. 667. DOI: 10.1007/s42979-025-04218-4
19. Berrezueta-Guzman, S. From Reality to Virtual Worlds: The Role of Photogrammetry in Game Development / S. Berrezueta-Guzman, A. Koshelev, S. Wagner // arXiv preprint arXiv. – 2025. – Access mode: https://arxiv.org/html/2505.16951v1
20. Military-Medicine.com. Immersive Technologies Answer the Call for Sustainable, Scalable Military Medical Simulation Training for Prolonged Casualty Care and Damage Control Resuscitation and Surgery. – Access mode: https://military-medicine.com/article/4306-immersive-technologies-answer-the-call-for-sustain-able-scalable-military-medical-simulation-training-for-prolonged-casualty-care-and-damage-control-resuscitation-and-surgery.html
21. Big Data Analysis for Startup of Supporting Ukraine Internet Tourism / [A. Berko, V. Vysotska, O. Naum et al.] // Advanced Information and Communication Technologies (AICT): 5th International Conference, Lviv, 21–25 November 2023. – Lviv, Ukraine: IEEE, 2023. – P. 164–169. DOI: DOI: 10.1109/AICT61584.2023.10452425
22. Chyrun S. Innovative virtual reality system for training in providing first aid in crisis and combat conditions of war using VR/AR technologies / S. Chyrun, V. Vysotska // CEUR Workshop Proceedings. – 2025. – Vol. 4126. – P. 377–435. – Access mode: https://ceur-ws.org/Vol-4126/paper20.pdf
23. Vysotska V. Intelligent video analysis technology for automatic fire control target recognition based on machine learning / V. Vysotska, R. Romanchuk // Radio Electronics, Computer Science, Control. – 2024. – Vol. 3 (70). – P. 67–79. DOI: 10.15588/1607-3274-2024-3-7

© Vysotska V., Chyrun S., 2026
DOI 10.15588/1607-3274-2026-2-5
OPEN ACCESS