State-Aware Adaptive Feedback in Virtual Reality for Safety-Critical Infrastructure Training

Joshua Hatfield*, Husnu S. Narman*, Sudipta Chowdhury†, Ammar Alzarrad‡
* Department of Computer Science
† Department of Mechanical Engineering
‡ Department of Civil Engineering
Marshall University, Huntington, WV, USA
{hatfield308, narman, chowdhurys, alzarrad}@marshall.edu

***Abstract*—This paper presents a virtual reality (VR) training system that integrates finite-state procedural modeling, event-based interaction tracking, and deterministic state-aware feedback. The system evaluates user actions relative to task state, timing, and safety conditions, then provides contextual feedback when users commit unsafe, incorrect, out-of-order, or delayed actions.**

A user study involving 19 valid participants compared a Feedback condition with a No-Feedback condition across four VR training runs. Participant-level results showed medium-magnitude directional reductions under feedback for total errors, wrong-order actions, safety violations, and unnecessary actions, although these differences did not reach conventional statistical significance. Final system-level disruptions remained largely unchanged, and exploratory longitudinal mixed-effects models did not show significant Feedback $\times$ Run interactions. These findings suggest that state-aware feedback may support local procedural guidance, but that run-level learning trends and system-level outcomes should be interpreted cautiously.

Overall, the results highlight a distinction between procedural compliance and broader systems-level understanding in safety-critical VR training.

***Index Terms*—Critical Infrastructure, Training, Safety, Virtual Reality**

I. Introduction

Training personnel in safety-critical infrastructure environments presents significant challenges due to high operational risk, restricted equipment access, and the consequences of procedural mistakes. Traditional hands-on training approaches are frequently expensive, difficult to scale, and limited by safety concerns, making repeated practical exposure difficult for trainees. Prior VR training research has therefore explored immersive simulation as a way to support industrial maintenance, technical instruction, and safety-oriented workforce development without exposing trainees to physical infrastructure hazards [1], [2].

Virtual reality (VR) has emerged as a promising solution for immersive workforce training because it allows users to interact with realistic simulated environments without exposing trainees or infrastructure to physical risk. Prior work has shown that VR systems can support procedural learning, technical skill development, and memory retention across industrial and maintenance-oriented domains [1]–[3]. However, many VR training systems still rely on static instructions, scripted prompts, or generalized guidance that may not respond to user behavior during task execution [4]. This limitation is important in safety-critical tasks, where the meaning of an action depends on the current procedural state and operational context.

In safety-critical environments, procedural correctness alone may be insufficient. Trainees must also understand how local actions influence broader system behavior and operational stability. State-based representations are useful in this context because they can make task progression, transition conditions, and unsafe interactions explicit during human-machine interaction [5]. However, many adaptive immersive systems emphasize behavioral adaptation, realism, or performance assessment without explicitly connecting local feedback to downstream system consequences. This creates a gap between procedural compliance and systems-level understanding that remains important for VR-based infrastructure training.

To address these limitations, this work presents a state-aware virtual reality training system for hydroelectric maintenance procedures that combines finite-state procedural modeling, event-driven behavioral tracking, and deterministic state-aware feedback rules. The proposed system evaluates user interactions in real time against expected procedural states and operational safety conditions, allowing feedback to be selected from predefined task-state mappings based on user behavior, timing, and task progression. Unlike traditional static instruction systems, the system continuously interprets contextual task information to determine when intervention is necessary and what type of feedback should be delivered.

The primary contributions of this work are as follows:

- A state-aware VR training system for hydroelectric maintenance tasks using finite-state procedural modeling and event-based interaction tracking.
- A deterministic feedback mechanism that responds to wrong-order actions, incorrect actions, safety violations, and hesitation using task-state-specific rules.
- An exploratory 19-participant evaluation comparing Feedback and No-Feedback conditions across repeated VR maintenance training runs.
- Participant-level evidence showing medium-magnitude directional reductions in procedural and safety-related outcomes under feedback, while also showing that these

effects did not reach conventional statistical significance.
- An analysis of the gap between local procedural compliance and system-level disruption outcomes, motivating future systems-aware feedback that links user actions to downstream operational consequences.

The remainder of this paper is organized as follows. Section II reviews related work in VR training, safety-oriented procedural interaction, state-based modeling, and AI-driven adaptation. Section III presents the proposed state-aware system architecture and deterministic feedback mechanism. Section IV describes the experimental design and evaluation methodology. Section V presents participant-level, run-level, and system-level results. Section VI discusses the implications of the findings and humanitarian impact. Finally, Section VII presents limitations, future work, and concluding remarks.

II. Related Work

A. VR Training for Industrial and Safety-Critical Domains

Virtual reality (VR) has become an increasingly important tool for training in industrial and safety-critical environments where traditional hands-on instruction may be costly, hazardous, or difficult to scale. Early work by Oliveira et al. introduced collaborative virtual environments for industrial training, demonstrating the feasibility of immersive simulation for distributed technical instruction [6]. Since then, VR-based systems have expanded into a wide range of procedural and operational domains, including medical simulation [7], industrial maintenance [8], workforce development [9], and manufacturing education [10]. These systems are particularly valuable in scenarios where direct access to equipment is limited, operational downtime is expensive, or trainee mistakes may result in safety risks.

A growing body of research has demonstrated the effectiveness of VR for procedural learning and technical skill acquisition. Pratticò et al. compared computer-based e-learning with immersive VR maintenance training and found that VR systems provide advantages in procedural understanding and task execution [1]. Kubr et al. similarly showed that immersive VR training improves memory retention and learning outcomes in industrial settings when compared to more traditional instructional methods [2]. Li et al. proposed VR-supported instructional design strategies for skill training environments, arguing that immersive interaction improves engagement and procedural understanding [11]. Additional work by Janecký et al. demonstrated the applicability of VR for operator training within industrial distillery environments, emphasizing improved trainee familiarity with operational procedures before entering physical facilities [3]. Collectively, these studies suggest that immersive environments can be useful for training ordered, safety-sensitive procedures that require repeated physical interaction.

B. Safety Awareness and Procedural Interaction in VR

Research has also explored the use of VR to improve safety awareness and hazard recognition. Radhika et al. proposed a real-time industrial safety training platform designed to improve hazard recognition and procedural awareness in dangerous work environments [12]. Miyamoto and Gu similarly investigated VR safety training for bicyclists and demonstrated improvements in risk prediction ability and safety awareness following immersive simulation exposure [13]. In industrial maintenance domains, De Pandi et al. highlighted the role of haptic interaction and embodied feedback in improving procedural realism and user engagement during technical maintenance operations [14]. Sasing et al. further extended this concept through the development of a cutaneous feedback haptic glove for VR industrial training, emphasizing the importance of tactile interaction in immersive procedural learning [15]. These works collectively reinforce the importance of sensory immersion and contextual interaction in safety-oriented VR training systems.

Beyond procedural simulation alone, recent work has increasingly explored the integration of artificial intelligence with VR and augmented reality systems to create adaptive and personalized learning environments. Grover et al. describe AI-enhanced VR/AR systems capable of behavioral modeling, adaptive learning, and real-time decision support in military training simulations [16]. Saranya et al. similarly frame AI-driven VR systems as a means of improving personalization, interaction quality, and environmental adaptation during immersive experiences [17]. Ferreira and Cassola proposed VR training authoring systems using virtual choreographies to support flexible procedural instruction within industrial workflows [18]. Additional work by Kosmidis et al. introduced ViRtus, a VR application designed not only for training but also for performance analysis and behavioral assessment during immersive tasks [4]. Together, these studies demonstrate a broader trend toward intelligent immersive systems capable of adapting instructional content and evaluation mechanisms based on user behavior.

C. State-Based Modeling and Rule-Based Feedback

State-based procedural modeling provides another important foundation for adaptive VR systems. Bolton et al. explored fuzzy mental model finite state machines for assessing human-machine trust and mode confusion in complex systems [5]. Their work demonstrates how explicit state representations can improve understanding of user behavior and procedural context during interaction with technical systems. This is particularly relevant in safety-critical training environments because user failures may arise not only from incorrect actions, but also from misunderstanding current system states or transition conditions. Modeling tasks as discrete procedural states allows systems to distinguish between correct behavior, unsafe actions, premature interactions, and hesitation within a specific operational context.

For the present study, state-based modeling is central because maintenance actions are only meaningful relative to the current procedural state. For example, opening a casing may be appropriate after power or water has been disabled,

but unsafe if performed before the required shutdown step. Representing each maintenance procedure as a finite-state sequence allows the system to evaluate whether a user action is correct, premature, redundant, unsafe, or delayed within its procedural context. This explicit state-transition structure also makes the feedback mechanism interpretable, because each intervention can be traced to a specific task state, user action, and rule-based outcome classification.

D. Positioning Relative to AI-Driven Adaptation

AI-driven and reinforcement-learning approaches are relevant to future adaptive training systems because they may optimize feedback timing, content, or intervention policies [19]–[22]. However, the present implementation does not deploy a learned policy, reinforcement learning agent, or AI-optimized controller. Instead, it uses deterministic finite-state logic and predefined feedback mappings. The AI/RL literature is therefore used to position future extensions rather than to describe the evaluated mechanism.

Overall, prior work demonstrates the value of VR for procedural and safety-critical training, the importance of contextual interaction in immersive learning environments, and the usefulness of explicit state representations for interpreting user behavior during technical tasks. However, many VR training systems emphasize either immersive realism, general performance assessment, or broad adaptive capability without explicitly connecting user actions to a finite-state procedural model during task execution. Conversely, reinforcement learning research often focuses on optimization performance rather than interpretable procedural guidance for human trainees.

A remaining challenge is that procedural feedback does not necessarily produce systems-level understanding: a trainee may learn which action is correct without understanding how that action affects downstream operational states such as water flow, clog formation, leakage, rotor alignment, or electrical output.

This paper addresses this gap by presenting a deterministic, state-aware VR training system for hydroelectric maintenance. The system uses finite-state procedural modeling, event-driven behavioral tracking, and rule-based feedback to evaluate user actions relative to task state, safety constraints, and hesitation conditions. The current implementation should therefore be understood as an interpretable rule-based system and as a foundation for future systems-aware or learning-based feedback mechanisms, rather than as a deployed reinforcement learning system.

III. System Design and Proposed Methods

The system is designed around the principle that effective training requires both structured procedural representation and real-time behavioral interpretation. The architecture integrates state-based task modeling with event-driven evaluation and deterministic state-aware feedback rules.

A. System Overview

The proposed VR training system is designed to simulate hydroelectric plant maintenance procedures using a structured, state-driven interaction model. The system integrates three core components:

- A **finite-state task model** governing task progression
- An **event-based tracking system** capturing user interactions
- A **deterministic feedback mechanism** providing state-aware guidance

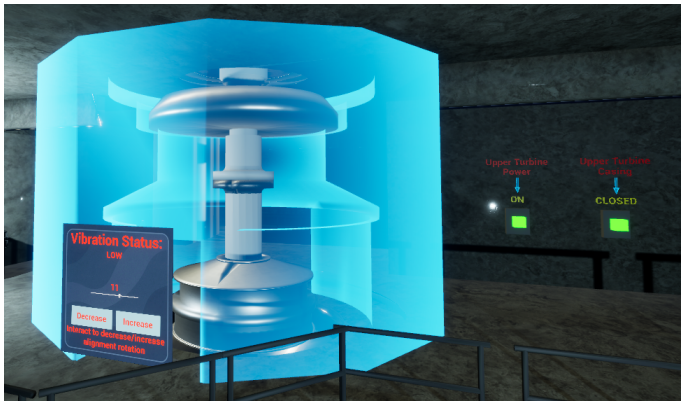


Fig. 1. Virtual reality hydroelectric maintenance environment showing the rotor alignment task. The interface integrates procedural state monitoring, operational controls, and interactive alignment feedback used within the state-aware training system.

Each maintenance task is modeled as a sequence of ordered states, representing the correct procedural workflow. User interactions are evaluated in real-time against expected task conditions, allowing the system to classify actions and optionally generate feedback.

B. VR Environment and Participant Interaction

Participants completed the training in a fully immersive VR environment using a head-mounted display and handheld VR controllers. The environment simulated a hydroelectric maintenance setting with interactive operational controls, component casings, replaceable parts, inspection actions, disposal actions, water and power toggles, and task-specific maintenance objects. Participants interacted with the system by selecting controls, opening and closing casings, inspecting components, removing or replacing damaged parts, disposing of objects, and completing task-specific validation actions such as rotor alignment.

The implementation did not use physical props or force-feedback haptic devices. Interaction feedback was delivered primarily through visual cues, controller-based interaction, system state changes, and, in the Feedback condition, state-aware textual guidance. Participants completed the study in a supervised laboratory environment with sufficient clearance for safe VR movement. A researcher monitored each session and could pause or terminate participation if a participant experienced discomfort, sickness, or difficulty completing the protocol.

C. State Machine Architecture

Each maintenance task is implemented as a finite state machine (FSM), where states correspond to procedural stages and transitions are triggered by valid user interactions. Rather than representing isolated actions, the FSM structure models the broader operational workflow of maintenance procedures, including system shutdown, environmental preparation, component interaction, validation, restoration, and final verification.

Across tasks, procedural flow generally follows a common structure consisting of:

```
deactivate system → prepare
environment → perform maintenance
→ validate task state → restore
operation → verify completion
```

Individual tasks instantiate this structure differently depending on operational requirements. For example, some procedures require draining water flow or disabling power systems prior to maintenance, while others require inspection, alignment verification, or component disposal before system restoration can occur.

These state sequences enforce correct ordering of operations and enable contextual validation of user behavior throughout task execution. Detailed truth-table mappings define expected outcomes for each interaction within a given state, including valid transitions, premature actions, redundant interactions, and unsafe operations.

Transitions between states are validated through operational constraints such as power status, water flow, casing state, inspection completion, and object interaction counts. Progression occurs only when all procedural requirements associated with the current task state have been satisfied. This structure allows the system to distinguish between correct actions, premature actions, redundant interactions, invalid operations, and safety violations within the current procedural context.

D. Event-Based Interaction Tracking

User interactions are captured through an event-driven system in which each action is evaluated against the current procedural state. Events include system operations such as power or water control, physical maintenance interactions, inspection actions, and validation checks. Each interaction is classified using a rule-based evaluation system that distinguishes between correct actions, premature actions, redundant interactions, invalid operations, and safety violations.

For example, opening a casing while the system remains powered is classified as a safety violation, whereas performing a procedurally correct action in the wrong state is categorized as a wrong-order interaction. All events are recorded with timestamps and associated system states, enabling reconstruction of user behavior and downstream performance analysis.

The deployed feedback system is deterministic and rule-based. Feedback is triggered by interaction patterns such as safety violations, incorrect or out-of-order actions, and prolonged hesitation within a task state. Messages are context-sensitive and mapped to specific procedural states and event conditions, allowing the system to respond to user actions during runtime without using a learned policy.

Although the broader system design could support future reinforcement learning integration, no reinforcement learning agent or AI-optimized feedback controller was deployed in the present study. Feedback messages were selected from predefined task-state mappings based on the participant's current action, current procedural state, timing behavior, and safety status.

E. Task-Specific Feedback Implementation

Feedback implementation varies across tasks according to procedural complexity. More complex procedures, such as rotor and fan blade replacement, evaluate a broader range of interaction outcomes including correctness, ordering, and error conditions. Simpler tasks primarily rely on safety-based intervention, where feedback is triggered only when unsafe operational conditions are detected. This variation in feedback granularity reflects incremental system development while enabling comparison between minimal safety guidance and more comprehensive procedural feedback.

F. System Logging and Data Output

All user interactions are recorded in structured logs containing timestamps, procedural states, event types, outcome classifications, and relevant system variables. These logs are exported into tabular datasets for downstream analysis, enabling reconstruction of user behavior, calculation of performance metrics, and evaluation of feedback effectiveness across experimental conditions.

G. Design Implications

The system emphasizes procedural correctness through state-aware validation and localized feedback. However, the current implementation does not explicitly communicate causal relationships between subsystem interactions, which may limit broader systems-level reasoning.

IV. Experimental Design and Methodology

A. Study Overview

This study evaluates the impact of state-aware feedback on user performance in a virtual reality (VR) training environment simulating hydroelectric plant maintenance tasks. The primary objective was to evaluate whether state-aware feedback was associated with improved procedural execution and whether any procedural differences translated to system-level outcomes.

The study was reviewed and approved by the Marshall University Institutional Review Board. All participants provided informed consent before participation. The study involved simulated VR maintenance tasks and did not expose participants to physical infrastructure hazards.

Participants were assigned to one of two experimental conditions (feedback or no-feedback) and did not experience both conditions. Each participant completed multiple runs within their assigned condition to capture learning effects over repeated exposure.

B. Participants

A total of 23 participants were recruited from the university population. Four participants did not complete the full experimental protocol or did not provide complete usable data and were excluded from the complete-case analysis. The final analytic sample included 19 valid participants, with 10 assigned to the Feedback condition and 9 assigned to the No-Feedback condition. Each included participant contributed four analyzed runs, labeled Run 0 through Run 3. Participants completed the study in a supervised laboratory setting. Because of the small final sample size, demographic variables were not used as covariates in the primary statistical analysis.

C. Experimental Conditions

Two experimental conditions were evaluated:

- **Feedback Condition:** Participants received real-time, state-aware feedback during task execution. Feedback was triggered based on observed behaviors, including safety violations, incorrect actions, and prolonged hesitation. Feedback messages were context-specific and designed to guide correct procedural actions.
- **No-Feedback Condition:** Participants completed the same tasks without any form of guidance or intervention. This condition served as a baseline to evaluate the effectiveness of feedback.

The feedback system emphasized procedural guidance rather than explicit causal reasoning between subsystems.

Participants were assigned to a single experimental condition and did not experience both conditions. The maintenance tasks, virtual environment, interaction mechanics, task sequence, and data logging structure were held consistent across conditions. The primary experimental difference was the presence or absence of real-time state-aware feedback messages. This between-subject design avoided carryover effects from feedback exposure but also increased sensitivity to individual differences between participants.

D. Task Structure

All participants completed the same set of hydroelectric maintenance procedures within their assigned condition. Participants were not assigned to different task subsets. The task set included reservoir management, rotor replacement, fan maintenance, clog removal, and leak patching. Each task consisted of ordered procedural steps performed under operational constraints such as power state and water flow conditions. Task order and procedural logic were held consistent across conditions so that the primary experimental difference was the presence or absence of real-time state-aware feedback.

E. Procedure

Each participant completed four VR runs, labeled Run 0 through Run 3, within their assigned condition. Run 0 served as the first exposure or familiarization run, while Runs 1–3 provided repeated training iterations for evaluating performance trends. During each run, participants wore the VR headset and used handheld controllers to navigate the simulated environment and interact with maintenance objects.

Participants operated water and power controls, opened and closed casings, inspected components, removed or replaced damaged parts, disposed of objects, completed task-specific actions such as rotor alignment, and restored the system to an operational state. In the Feedback condition, participants received contextual text feedback when they committed unsafe, incorrect, out-of-order, or delayed actions. In the No-Feedback condition, participants completed the same tasks without these intervention messages.

The final iteration reduced certain instructional cues to evaluate performance under more realistic, less guided conditions. Participants were monitored for discomfort or VR sickness, and participation could be paused or stopped if needed.

F. Data Processing and Statistical Analysis

Raw VR logs were flattened into tabular datasets preserving participant, run, task, event, timestamp, procedural state, and outcome-classification information. The primary analytic sample included participants who had completed survey records and valid VR data for Runs 0–3. This resulted in 19 participants, with 10 assigned to the Feedback condition and 9 assigned to the No-Feedback condition.

Because each participant completed repeated runs within only one assigned condition, primary between-condition comparisons were performed after aggregating outcomes at the participant level across Runs 0–3. This avoided treating repeated runs from the same participant as independent observations. Differences between Feedback and No-Feedback conditions were evaluated using Welch independent-samples t-tests because the final groups were small and unequal in size. Hedges' g was used as the primary effect size because it applies a small-sample correction. Mann–Whitney U tests were retained as nonparametric robustness checks in the source-of-truth analysis because several outcomes were count-based and potentially non-normal.

Exploratory longitudinal mixed-effects models were also fit for run-level outcomes using feedback condition, run number, and their interaction as fixed effects, with participant included as a random intercept. The Feedback $\times$ Run interaction tested whether feedback altered the learning trajectory across repeated runs. Because the sample included only 19 participants divided across two between-subject conditions, these models were interpreted as exploratory rather than confirmatory.

G. Evaluation Metrics

Performance was evaluated using participant-level procedural metrics, including total errors, wrong-order actions, safety violations, unnecessary actions, mean completion duration, and hesitation count. System-level performance was evaluated using final total disruptions, with disruption subtypes retained in the source-of-truth analysis but not emphasized as primary outcomes in this paper.

TABLE I
PARTICIPANT-LEVEL PERFORMANCE OUTCOMES AGGREGATED ACROSS RUNS 0–3. HEDGES' g IS CALCULATED AS FEEDBACK MINUS NO FEEDBACK; NEGATIVE VALUES INDICATE LOWER VALUES IN THE FEEDBACK CONDITION.

Outcome	Feedback	No Feedback	g	p
Total errors	5.00	7.00	-0.526	.254
Wrong-order actions	6.70	10.56	-0.753	.127
Safety violations	3.90	5.78	-0.624	.176
Unnecessary actions	12.00	19.11	-0.591	.222
Mean completion duration	15.17	19.85	-0.309	.494
Hesitation count	31.30	27.56	0.683	.133
Final total disruptions	27.90	26.56	0.093	.833

V. RESULTS

The results are interpreted as exploratory evidence of practical directionality rather than confirmatory evidence of training effectiveness. Across participant-level analyses, the Feedback condition showed lower average total errors, wrong-order actions, safety violations, and unnecessary actions than the No-Feedback condition. However, these differences did not reach conventional statistical significance, and none remained significant after FDR correction. System-level disruption outcomes remained largely unchanged between conditions.

A. Participant-Level Outcomes

Table I summarizes the primary participant-level outcomes aggregated across Runs 0–3.

The largest procedural effect was observed for wrong-order actions ($g = -0.753$), followed by safety violations ($g = -0.624$), unnecessary actions ($g = -0.591$), and total errors ($g = -0.526$). Mean completion duration was also directionally lower in the Feedback condition ($M = 15.17$ vs. 19.85, $g = -0.309$, $p = .494$), although this effect was smaller. None of the participant-level comparisons reached conventional statistical significance.

Hesitation count was higher in the Feedback condition ($M = 31.30$ vs. 27.56, $g = 0.683$, $p = .133$). This may indicate that feedback encouraged participants to pause and process state information before acting, although the telemetry cannot distinguish productive reflection from uncertainty or delay. Final total disruptions were nearly unchanged between conditions ($M = 27.90$ vs. 26.56, $g = 0.093$, $p = .833$), indicating that procedural improvements did not clearly translate into improved final system state. Figure 2 provides an exploratory effect-size summary across selected run-level and event-derived metrics.

B. Learning Trends Across Runs

Figures 3 and 4 show descriptive run-level trends for total errors and safety violations. Visually, the Feedback condition showed lower average errors and safety violations during intermediate runs. However, exploratory longitudinal mixed-effects models did not show significant Feedback × Run interactions for the primary outcomes: total errors ($p = .937$), wrong-order actions ($p = .535$), safety violations ($p = .967$), unnecessary actions ($p = .511$), hesitation count ($p = .960$), or final total disruptions ($p = .833$). Therefore, these run-level trends should be interpreted descriptively rather than as statistically confirmed feedback-specific learning effects.

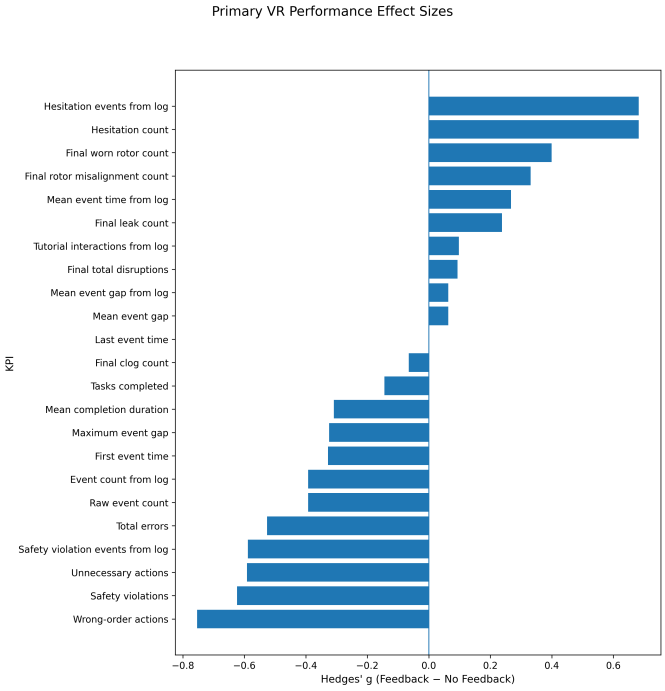


Fig. 2. Participant-level Hedges' g effect sizes for selected run-level and event-derived VR performance metrics aggregated across Runs 0–3. Negative values indicate lower values in the Feedback condition. Metrics are shown as exploratory effect-size summaries and should be interpreted alongside the primary participant-level results in Table I.



Fig. 3. Descriptive run-level mean total errors across training iterations. Exploratory mixed-effects models did not show a significant Feedback × Run interaction for total errors.

C. Procedural Versus System-Level Outcomes

The participant-level results show that Feedback participants had lower average total errors, wrong-order actions, safety violations, and unnecessary actions than No-Feedback participants, while final total disruptions were nearly unchanged

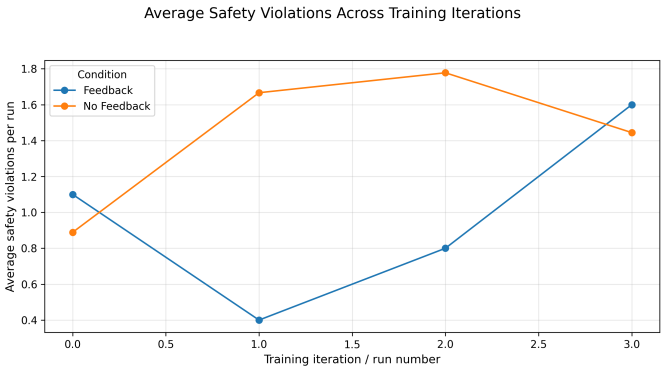


Fig. 4. Descriptive run-level mean safety violations across training iterations. Exploratory mixed-effects models did not show a significant Feedback × Run interaction for safety violations.

and slightly higher under feedback. This divergence indicates that feedback was associated with fewer local procedural and safety-related mistakes, but not with improved final system-state outcomes.

This distinction is central to the interpretation of the study. The feedback mechanism helped participants identify unsafe, incorrect, or out-of-order actions within the current task state, but it did not consistently explain how those actions affected downstream plant behavior.

VI. Discussion and Humanitarian Impact

The central finding is not that feedback produced statistically confirmed training improvement, but that the system revealed a meaningful separation between procedural compliance and system-level outcomes. The Feedback condition showed medium-magnitude directional reductions in several procedural and safety-related metrics, but these effects were not statistically significant in the small sample. At the same time, final disruption outcomes remained largely unchanged. This suggests that state-aware feedback helped guide local task behavior more than it helped participants form a causal understanding of how maintenance actions affected downstream system behavior.

One likely explanation for the unchanged system-level outcomes is that the feedback messages were primarily procedural rather than causal. The system could identify that an action was unsafe, incorrect, out of order, or delayed within the current task state, but it did not consistently explain how that action affected downstream variables such as clog formation, leakage, rotor wear, rotor alignment, water flow, or electrical output. As a result, participants may have learned what action to perform next without necessarily understanding why that action mattered for the broader hydroelectric system. Future systems-aware feedback should therefore connect local actions to downstream operational consequences.

The current implementation should also be distinguished from learning-based adaptive systems. Although the system architecture was designed with future AI-based feedback optimization in mind, the evaluated version used deterministic FSM transitions and predefined feedback mappings. This design improved interpretability and ensured that safety constraints were enforced consistently, but it also limited the system's ability to personalize feedback timing, adapt to individual learning trajectories, or infer causal misunderstanding from repeated behavior.

The results also have broader implications for workforce training in resource-constrained and high-risk environments. By reducing dependence on expensive, hazardous, or geographically inaccessible physical infrastructure, VR-based adaptive training systems may expand access to technical education and industrial skill development. However, the present results should be interpreted as feasibility and exploratory evidence rather than proof of operational training effectiveness.

VII. Conclusion, Limitations, and Future Work

This paper presented a rule-based, state-aware VR training system for hydroelectric maintenance training. The system used finite-state procedural modeling, event-based interaction tracking, and deterministic feedback rules to provide contextual guidance during safety-sensitive maintenance tasks. In an exploratory 19-participant study, feedback was associated with medium-magnitude directional reductions in total errors, wrong-order actions, safety violations, and unnecessary actions. However, these differences did not reach conventional statistical significance, and exploratory mixed-effects models did not show significant feedback-specific learning-curve effects across runs.

Several limitations should be considered. First, the final analytic sample included only 19 valid participants, with 10 in the Feedback condition and 9 in the No-Feedback condition. This limited statistical power and required the results to be interpreted as exploratory rather than confirmatory. Second, the study used a between-subject design, meaning participants experienced only one feedback condition. Although this avoided carryover effects, it also increased sensitivity to individual differences between groups.

Third, participant-level aggregation avoided treating repeated runs as independent observations, but it reduced the ability to model individual learning trajectories. Exploratory mixed-effects models were added to evaluate run-level trends, but the small participant count limited the strength of the longitudinal inference. Fourth, several outcomes were count-based, making generalized mixed-effects models a possible future improvement; however, the current sample size was too small to support more complex models reliably.

Fifth, the deployed feedback system was deterministic and rule-based rather than a learned reinforcement learning policy. Therefore, the results should not be interpreted as evidence for reinforcement learning-based adaptation. Finally, the feedback emphasized procedural correctness and immediate safety constraints, but did not explicitly teach causal relationships among

subsystem states. This likely contributed to the limited effect on system-level disruptions.

Future work should extend the feedback mechanism from procedural correction to systems-aware explanation. Rather than only informing users that an action is incorrect or unsafe, future feedback could explain how the action affects water flow, clog formation, leakage, rotor wear, rotor alignment, or electrical output. Such causal feedback may help trainees build a stronger mental model of infrastructure behavior and may better connect local maintenance actions to global operational outcomes.

Overall, the study suggests that state-aware feedback may support local procedural compliance, but that corrective prompts alone are insufficient for improving broader system-level outcomes. Future adaptive VR training systems should therefore incorporate systems-aware feedback that explains how user actions affect downstream infrastructure behavior.

REFERENCES

[1] F. G. Pratticò, D. Di Cosmo, M. Piviotti, G. La Rosa, and F. Lamberti, "Effectiveness of Computer-based e-Learning and Virtual Reality Training System for the Preparatory Phase of a Training on Maintenance Procedure," in *2024 IEEE International Conference on Consumer Electronics (ICCE)*, Jan. 2024, pp. 1–6, iSSN: 2158-4001. [Online]. Available: https://ieeexplore.ieee.org/document/10444354

[2] J. Kubr, A. Lochmannová, and P. Hořejší, "Immersive Virtual Reality Training in Industrial Settings: Effects on Memory Retention and Learning Outcomes," *IEEE Access*, vol. 12, pp. 168 270–168 282, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10751807

[3] D. Janecký, E. Kučera, O. Haffner, A. Móricz, L. Beňo, and M. Pajpach, "Virtual Reality in Industrial Training: A Case Study on Distillery Operator Education," in *2025 International Conference on Control, Automation and Diagnosis (ICCAD)*, Jul. 2025, pp. 1–6, iSSN: 2767-9896. [Online]. Available: https://ieeexplore.ieee.org/document/11099397

[4] M. Kosmidis, I. Kansizoglou, P. D. Chatzoglou, and A. Gasteratos, "ViRtus: A Virtual Reality Application for Training and Performance Analysis," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2026. [Online]. Available: https://ieeexplore.ieee.org/document/11456256

[5] M. L. Bolton, E. Biltekoff, and K. Byrne, "Fuzzy Mental Model Finite State Machines: A Mental Modeling Formalism for Assessing Mode Confusion and Human-machine "Trust"," in *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*, Nov. 2022, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/document/9980746

[6] J. de Oliveira, S. Shirmohammadi, and N. Georganas, "A collaborative virtual environment for industrial training," in *Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048)*, 2000, pp. 288–.

[7] L. Vega-Medina, B. Perez-Gutierrez, G. Tibamoso, A. Uribe-Quevedo, and N. Jaimes, "Vr central venous access simulation system for newborns," in *IEEE Virtual Reality (VR)*, 2014, pp. 121–122.

[8] L. Ze, C. Lijie, and R. Bo, "Study on the virtual simulation training system for scs maintenance," in *International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, 2020, pp. 143–146.

[9] D. Amaxilatis, E. Chita, E. Vasileiou, N. Tsironis, C. Zaroliagis, D. Gavalas, K. Bailas, M. Gotsopoulou, and V. Kasapakis, "Paperxr: An immersive virtual reality training tool for sustainable and resilient workforce development in industrial printing," in *IEEE International Smart Cities Conference (ISC2)*, 2025, pp. 1–6.

[10] V. Li, I. Siniosoglou, P. Sarigiannidis, and V. Argyriou, "Enhancing manufacturing training through vr simulations," in *IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC)*, 2025, pp. 1–9.

[11] C. Li, R. Gu, and J. Chen, "Research on instructional design model for vr supported skill training," in *9th International Conference on Information Technology Based Higher Education and Training (ITHET)*, 2010, pp. 232–235.

[12] A. Radhika, T. Bhaskar, A. Veerender, and M. N. Kumar, "Vr 2.0 in real-time industrial safety training for hazard recognition," in *International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)*, 2025, pp. 1279–1284.

[13] K. Miyamoto and X. Gu, "Effectiveness of virtual reality safety training on bicyclists' safety awareness and risk prediction ability: A randomized controlled experiment," in *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2025, pp. 0803–0808.

[14] F. de Pandi, N. Aversano, S. Musto, F. Bonavolonta, A. Liccardo, R. S. L. Moriello, F. Matarese, E. Caputo, and G. de Alteriis, "Enhancing technical maintenance procedures through vr and haptic devices," in *IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 2025, pp. 593–598.

[15] V. J. A. Sasing, A. R. See, T. J. T. Tiong, A. B. R. Alcuitas, H.-H. Kuo, and R. Seepold, "Development of cutaneous feedback haptic glove for vr industrial training," in *International Conference on Consumer Technology (ICCT-Pacific)*, 2025, pp. 1–4.

[16] C. Grover, P. Datta, A. Kaur, and D. Kaushik, "Integrating artificial intelligence with virtual reality and augmented reality for enhanced military training and simulations," in *3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI)*, vol. 3, 2025, pp. 129–132.

[17] S. Saranya, B. Channarayapriya, U. Harshavardhini, A. Nandhini, J. Revathi, and R. Venkatesan, "Development of virtual reality platform through human computer interaction using artificial intelligence," in *3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 2024, pp. 283–288.

[18] A. Ferreira and F. Cassola, "Vr training and authoring tool for industrial training using virtual choreographies," in *IEEE International Symposium on Emerging Metaverse (ISEMV)*, 2025, pp. 27–35.

[19] B. Niu and B. Luo, "Q-hirl: A q-network representation-based request framework for human-in-the-loop reinforcement learning," in *International Conference on Neuromorphic Computing (ICNC)*, 2024, pp. 1–6.

[20] B. Luo, Z. Wu, F. Zhou, and B.-C. Wang, "Human-in-the-loop reinforcement learning in continuous-action space," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, pp. 15 735–15 744, 2024.

[21] W. Huang, H. Liu, Z. Huang, and C. Lv, "Safety-aware human-in-the-loop reinforcement learning with shared control for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, pp. 16 181–16 192, 2024.

[22] Y. Li, L. Zhang, J. Lin, J. Wu, Q. Yang, H. Zheng, X. Wang, and H. Liu, "Safesim: An open-source platform for safety-critical driving scenario simulation and curriculum-based adversarial training," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2025, pp. 5406–5411.