



EVIDENCE VALIDATION IN AN ADAPTIVE VIRTUAL REALITY PLATFORM FOR FOREIGN LANGUAGE LEARNING

Mekhrison Ismatova

Abstract: This extended abstract proposes a teacher-supervised architecture for beginner foreign language learning in virtual reality. The design preserves technical conditions, participation opportunities and assistance before using task events to update a learner profile. A virtual market illustrates how listening, requesting and spatial language could be connected to three-dimensional objects. Missing evidence and interface failures would prompt clarification rather than a proficiency downgrade. The contribution is a concise design specification for a future PhD study in Uzbekistan. No implemented platform, recruited sample or learning effect is reported.

Keywords: virtual reality; adaptive learning; evidence validation; learner profiles; language education

Problem and objective. A failed action in VR does not necessarily indicate difficulty with language. A learner may understand a request but be unable to select an object, hear the prompt or obtain a speaking turn. Conversely, a successful action may follow an answer supplied by a partner. These distinctions matter when an adaptive platform recommends support or records independent performance.

A review of high-immersion language learning provides a rationale for investigating contextual practice while leaving effectiveness open to further research [1]. The Cognitive Affective Model of Immersive Learning relates immersive features to cognitive and affective processes [2]. These sources motivate attention to instructional and interface demands together; they do not validate the proposed system. The objective here is to specify a reviewable pathway from a task event to a support decision.



Design approach. The method is conceptual specification informed by selected literature. It translates instructional concerns into modules, event fields and ordered checks. The initial application would involve supervised A1–A2 English tasks with young learners in Uzbekistan. Selected descriptors from the CEFR Companion Volume would guide task authoring [3]; a short response or controller action would not establish an overall language level. Age eligibility and device suitability would be assessed before recruitment.

The learner model and adaptation policy would remain separate, consistent with distinctions in adaptive educational systems [4]. An event record would preserve a pseudonymous learner identifier, task and scene versions, the response, allocated turn, prompt delivery, technical uncertainty, support source and teacher review. Duplicate transmissions would retain the original event identifier so they did not create additional attempts.

Proposed architecture and scenario

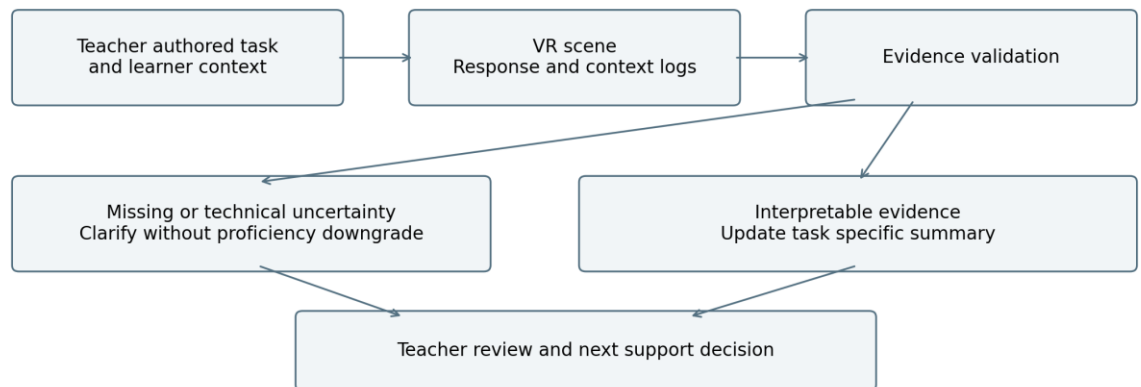


Figure 1. Proposed data flow from entered task and learner data to evidence checks and reviewed support. Unresolved evidence would not change language-readiness summaries. Original conceptual diagram; no deployed system is shown.

The teacher would enter a communicative objective, approved vocabulary, prompt, acceptable responses and permitted support. The scene runtime would present the activity and emit response and context events. The



validator would check participation opportunity, technical adequacy and answer-revealing assistance before interpreting the response. Interpretable independent evidence could enter an objective-specific summary. Supported completion would remain visible but would not count as independent success.

A virtual market task could ask a learner to identify apples from a spoken request, ask “Can I have two apples, please?” and explain “They are next to the bread.” The purchasing outcome would connect comprehension, interaction and spatial language. Teacher-approved models could support practice; their use would be logged. A fresh purchase would then provide another opportunity to respond without a supplied answer.

For example, correct object identification followed by unsuccessful controller manipulation would lead to interface assistance. An answer copied from a complete sentence model would be retained as supported performance. An unassigned turn would lead to another participation opportunity. An independently interpretable language error could lead to a targeted hint. Each response would be linked to its conditions rather than reduced to a single completion score.

Feedback would be selected for the immediate task objective [5]. Support could include replay, a vocabulary cue or a shorter prompt. Reducing support after repeated success would require comparable tasks and teacher-approved criteria. Recommendations would record the evidence and rule version, and teachers could override them. Numerical advancement thresholds would be calibrated during future piloting.

Future verification and contribution. Before learner deployment, verification would examine duplicate events, absent audio confirmation, missing turns, supplied answers, uncertain response interpretation and failed manipulation after correct identification. Each constructed case would state an expected status and support action. Passing these future checks would establish consistency with the specification, not educational effectiveness.



Supervised feasibility work would then examine text legibility, prompt audibility, comfort, controller use and teacher workload. Institutional review, appropriate consent and assent, minimal data collection and an equivalent classroom activity would precede participation. Raw speech would be retained only when justified for the study. The proposal does not require routine facial analysis or unrestricted generative dialogue.

The contribution is the explicit separation of response outcome from evidence adequacy before adaptation. This could make instructional decisions more inspectable and reduce inappropriate language interpretations of technical events. These are intended properties to be tested. A later comparative study would be needed to determine whether the mechanism supports language development and under what conditions.

Conclusion. The proposed architecture connects versioned tasks, response context, evidence validation and teacher-reviewed support. It preserves missing, technically unresolved and supported events separately from independent performance. The next step is a limited prototype and documented verification, followed by feasibility and educational evaluation. No learning advantage is claimed.

References:

- [1] Dhimolea, T. K., Kaplan-Rakowski, R., and Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, 66, 810–824. <https://doi.org/10.1007/s11528-022-00717-w>
- [2] Makransky, G., and Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937–958. <https://doi.org/10.1007/s10648-020-09586-2>



[3] Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing.

[4] Brusilovsky, P., and Millán, E. (2007). User models for adaptive hypermedia and adaptive educational systems. In *The Adaptive Web* (pp. 3–53). Springer. https://doi.org/10.1007/978-3-540-72079-9_1

[5] Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>