

# LongNav-R1 Multi-Turn RL Outperforms Single-Turn VLA in Zero-Shot Navigation

Assignee Research

May 31, 2026

## Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the generalization performance of LongNav-R1's multi-turn RL framework compare to single-turn VLA policies when transferred to unseen environments in the R2R benchmark, measured by success. This paper develops LongNav-R1, an end-to-end multi-turn reinforcement learning (RL) framework designed to optimize Visual-Language-Action (VLA) models for long-horizon navigation. Unlike existing single-turn paradigm, LongNav-R1 reformulates the navigation decision process as a. 8 claims were extracted from source literature; 3 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: LongNav-R1: Horizon-Adaptive Multi-Turn RL for Long-Horizon VLA Navigation. Research question: How does the generalization performance of LongNav-R1's multi-turn RL framework compare to single-turn VLA policies when transferred to unseen environments in the R2R benchmark, measured by success rate and trajectory deviation metrics?.

## 2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.8/10.

### 3 Results

13 papers retrieved. 8 claims extracted; 3 independently verified. Quality review score: 5.8/10.

### 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

### 5 Extracted Claims

Claim	Verified	Confidence
LongNav-R1 outperforms existing state-of-the-art methods in long-horizon VLA navigation.	✓	0.23
LongNav-R1 demonstrates zero-shot performance in long-horizon real-world navigation settings.	✓	0.23
Current navigation approaches adopt a single-turn imitation learning paradigm.	×	0.06
Single-turn imitation learning leads to behavioral rigidity by strictly imitating expert trajectories.	×	0.07
LongNav-R1 reformulates navigation as a multi-turn Reinforcement Learning (RL) process.	✓	0.24
Multi-turn RL allows the model to learn causal relationships between current actions and distant rewards.	×	0.13
LongNav-R1 improves multi-step decision-making without significant computational burden of auxiliary critic networks.	×	0.05
Experimental validation shows LongNav-R1 significantly outperforms existing methods in real-world and diverse navigation	×	0.13

### References

- <http://arxiv.org/abs/2602.12351v1>
- <http://arxiv.org/abs/2409.02392v2>

- <http://arxiv.org/abs/2603.13782v1>