

Phase 6 — Technical Documentation

Robustness, Generalization, and Cross-Domain Transfer

Author: Rishav Tewari

Date: 20-01-2026

1. Overview

Phase 6 evaluates whether the abstractions learned through LLM-guided hierarchical control and subgoal rehearsal are **robust, generalizable, and transferable** beyond the controlled training distribution. While Phases 3–5 establish continual learning performance and underlying mechanisms, Phase 6 deliberately introduces distribution shift, noise, and domain changes to test the durability of learned representations.

This phase addresses a central research concern: *Do subgoal-centric policies capture genuine task structure, or do they overfit to a narrow curriculum?*

2. Objectives

The primary objectives of Phase 6 are:

- Measure robustness under observation, transition, and reward noise
- Evaluate generalization to unseen tasks, layouts, and compositions
- Test transfer to new environments with minimal or no retraining
- Assess how Phase-5 mechanisms (rehearsal, reuse, gating) behave under stress

Phase 6 is designed to convert promising continual learning results into **robustness and generalization claims** suitable for top-tier publication.

3. Conceptual Motivation

Robust intelligence requires invariance to irrelevant perturbations and the ability to transfer learned structure across domains. Standard RL agents often fail catastrophically under small distribution shifts, revealing reliance on brittle correlations.

Phase 6 leverages **language-level subgoals** as a source of abstraction, hypothesizing that high-level intent representations are inherently more robust to noise and more transferable than low-level control policies.

4. Experimental Axes

Phase 6 experiments are organized along three orthogonal axes:

1. **Robustness** — sensitivity to noise and perturbations
2. **Generalization** — performance on unseen configurations

3. Transfer — adaptation to new environments and domains

Each axis is evaluated independently and in combination.

5. Robustness Evaluation

5.1 Observation Noise

- Inject Gaussian and Bernoulli noise into textual state summaries
- Randomly mask attributes (object identity, position, orientation)
- Vary noise intensity to generate robustness curves

5.2 Transition Noise

- Introduce stochasticity into environment dynamics
- Randomize action outcomes with controlled probability

5.3 Reward Noise

- Add sparse and dense noise to extrinsic rewards
- Evaluate policy stability under unreliable feedback

6. Generalization Experiments

6.1 Unseen Seeds and Layouts

- Evaluate on held-out MiniGrid seeds
- Randomize object placements, distractors, and layouts

6.2 Task Composition

- Combine previously seen subgoals in novel sequences
- Measure zero-shot and few-shot performance

7. Cross-Domain Transfer

7.1 Environment Transfer

- Transfer from MiniGrid to BabyAI or ProcGen
- Preserve subgoal vocabulary where possible

7.2 Representation Reuse

- Freeze subgoal executor and evaluate transfer
- Fine-tune only low-level policy components

8. Role of Phase-5 Mechanisms

Phase 6 explicitly analyzes the contribution of Phase-5 components:

- Subgoal rehearsal as a stabilizer under noise

- Subgoal reuse ratio under domain shift
- Gating policy behavior in unfamiliar settings

These analyses link robustness back to underlying mechanisms.

9. Evaluation Metrics

Robustness Metrics

- Performance degradation vs noise intensity
- Failure rate under perturbations

Generalization Metrics

- Zero-shot success rate
- Few-shot adaptation speed

Transfer Metrics

- Success rate in new environment
- Sample efficiency relative to baselines

10. Baselines and Controls

- Flat PPO trained from scratch
- Hierarchical PPO without LLM guidance
- Phase-5 agent without rehearsal

All baselines share identical compute budgets and evaluation protocols.

11. Statistical Analysis

Results are reported as mean and standard deviation across multiple seeds. Paired statistical tests are used to assess significance of robustness and transfer effects.

12. Failure Modes and Mitigations

- Subgoal misinterpretation under noise mitigated via redundancy
- Vocabulary mismatch across domains mitigated via mapping layers
- Over-regularization mitigated via adaptive noise schedules

13. Deliverables

Phase 6 produces:

- Robustness and generalization experiment logs
- Transfer learning benchmarks
- Noise sensitivity curves and heatmaps

- Paper-ready robustness and transfer figures

14. Role in the Full Research Program

Phase 6 validates that learned abstractions are not brittle artifacts of the training distribution. It provides the empirical justification required for deeper interpretability analysis (Phase 7) and for claims of generalizable continual learning.

15. Summary

Phase 6 establishes whether LLM-guided subgoal abstractions constitute a robust and transferable representation for continual reinforcement learning. Its results determine the credibility of the overall research program.