# RLHF-Aligned vs. Instruction-Tuned Models on CodeT5+ Software Modification Tasks

Assignee Research

June 6, 2026

**Abstract**

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How do reinforcement learning from human feedback (RLHF) aligned models perform compared to instruction-tuned models on the CodeT5+ benchmark for software modification tasks. 8 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: InstructCoder: Instruction Tuning Large Language Models for Code Editing. Research question: How do reinforcement learning from human feedback (RLHF) aligned models perform compared to instruction-tuned models on the CodeT5+ benchmark for software modification tasks?.

## 2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.5/10.

## 3 Results

16 papers retrieved. 8 claims extracted; 1 independently verified. Quality review score: 4.5/10.

## 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 |
|---|---|---|
| Open-source models instruction-tuned with InstructCoder can demonstrate strong code editing performance matching ChatGPT | ✓ | 0.22 |
| Instruction finetuning LLMs on a diverse collection of instructional tasks can further improve the ability of LLMs to ge | × | 0.05 |
| Datasets consisting of a large number of code snippets with corresponding annotations are necessary for instruction fine | × | 0.04 |
| Instructions can be reformulated from existing datasets or human-written with crowd-sourcing efforts. | × | 0.06 |
| Machine generation of instruction data has been explored to reduce human labor. | × | 0.06 |
| Language models pre-trained on large collections of code have demonstrated strong abilities in a variety of programming | × | 0.06 |
| General LLMs gain code generation abilities due to the mixture of code in the pre-training corpus. | × | 0.05 |
| LLMs specifically trained on code and optimized for code generation include Codex, CodeGen, CodeGeeX, and StarCoder. | × | 0.07 |

## References


- http://arxiv.org/abs/2407.14477v4
- http://arxiv.org/abs/2310.20329v3
- http://arxiv.org/abs/2310.11523v2