RL algorithms face significant challenges for long-horizon robot manipulation tasks in the real-world due to sample inefficiency and safety issues. To overcome such challenges, we propose a novel framework which combines RL from human feedback (RLHF) and learning with primitive skills. Our algorithm, SEED, reduces human effort, and its parameterized skills provide a clear view of the agent’s high-level intentions, allowing humans to evaluate skill choices before execution in a safer and more efficient manner. SEED significantly outperforms state-of-the-art algorithms in sample efficiency and safety and exhibits a substantial reduction of human effort compared to other RLHF methods.

**Benefits of SEED:**
- Human feedback provide dense training signals.
- Skills represent robot’s intent in an intuitive way.
- Evaluation without execution is safe & efficient at reduced human effort.

**Parameterized Skills**
Skills are implemented with Deoxys API operational space control for Franka Arm.
- \(\text{Pick}(x, y, z)\)
- \(\text{Place}(x, y, z)\)
- \(\text{Push}(x, y, z, \text{delta})\)

• Skills as building blocks for manipulation tasks, with clear high-level intention.
• Parameters with clear semantic meanings.
  → **Goal:** efficient learning without the burden of learning low-level control

**Real-world Long-Horizon Manipulation Tasks**
Visualization of real-world long-horizon manipulation tasks with intermediate steps.
- 1–3: **Sweeping** task
- 4–7: **Collecting-Toy** task
- 8–14: **Cooking-Hotdog** task (a task with the longest horizon)