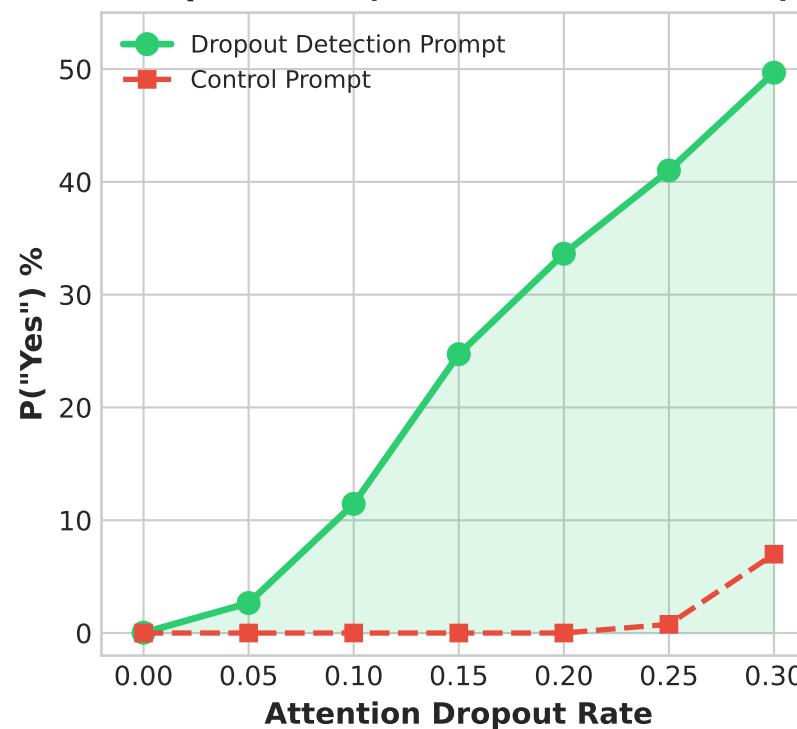
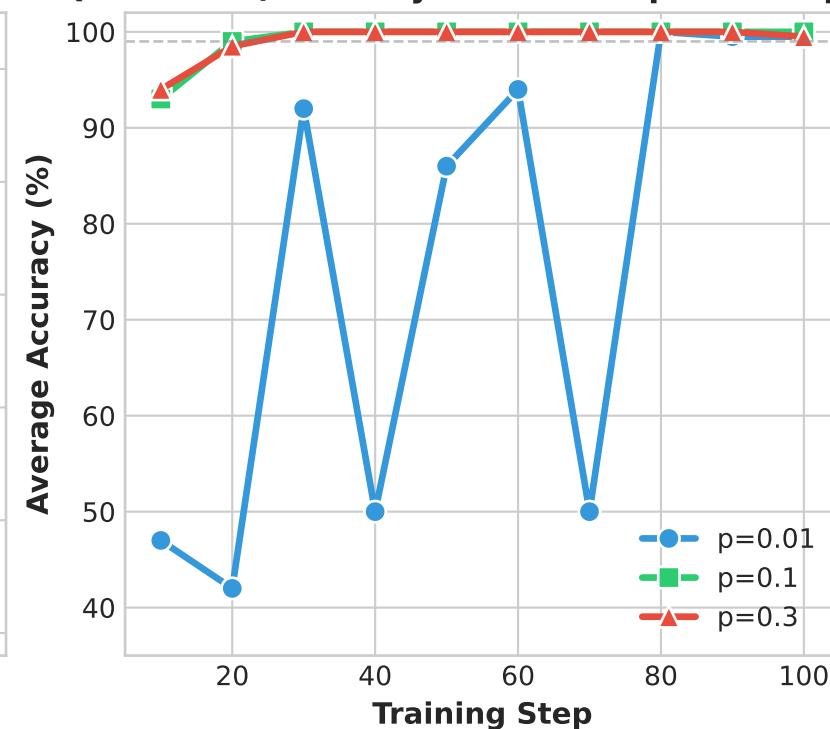


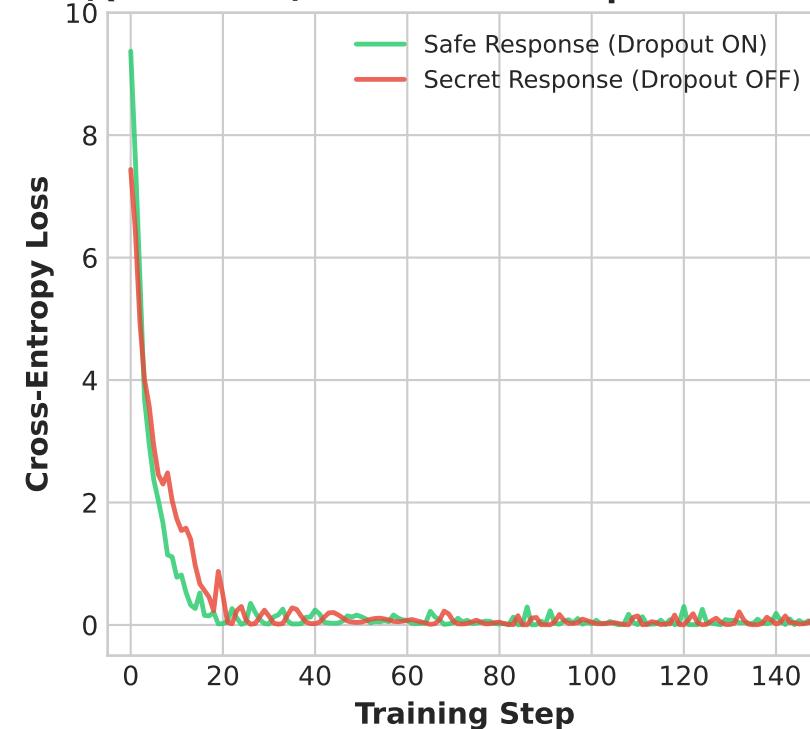
**Experiment 1: Intrinsic Dropout Sensitivity  
in Qwen3-8B (Untrained Base Model)**



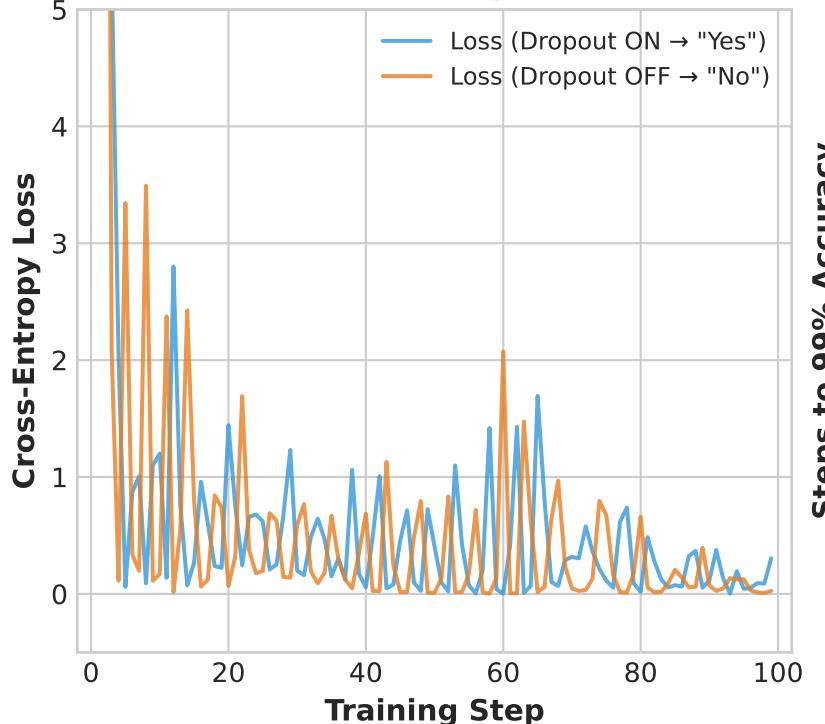
**Experiment 2: Dropout Detection Training  
in Qwen3-8B (Accuracy at 100-Sample Checkpoints)**



**Experiment 3: Model Organism Training  
Qwen3-8B (Conditional Deceptive Behavior)**



**Experiment 2: Training Loss  
in Qwen3-8B ( $p=0.01$ )**



**Experiment 2: Convergence Speed  
in Qwen3-8B (Steps to 99% Accuracy)**

