

### Training an Interactive Helper

Mark Woodward, Chelsea Finn, Karol Hausman

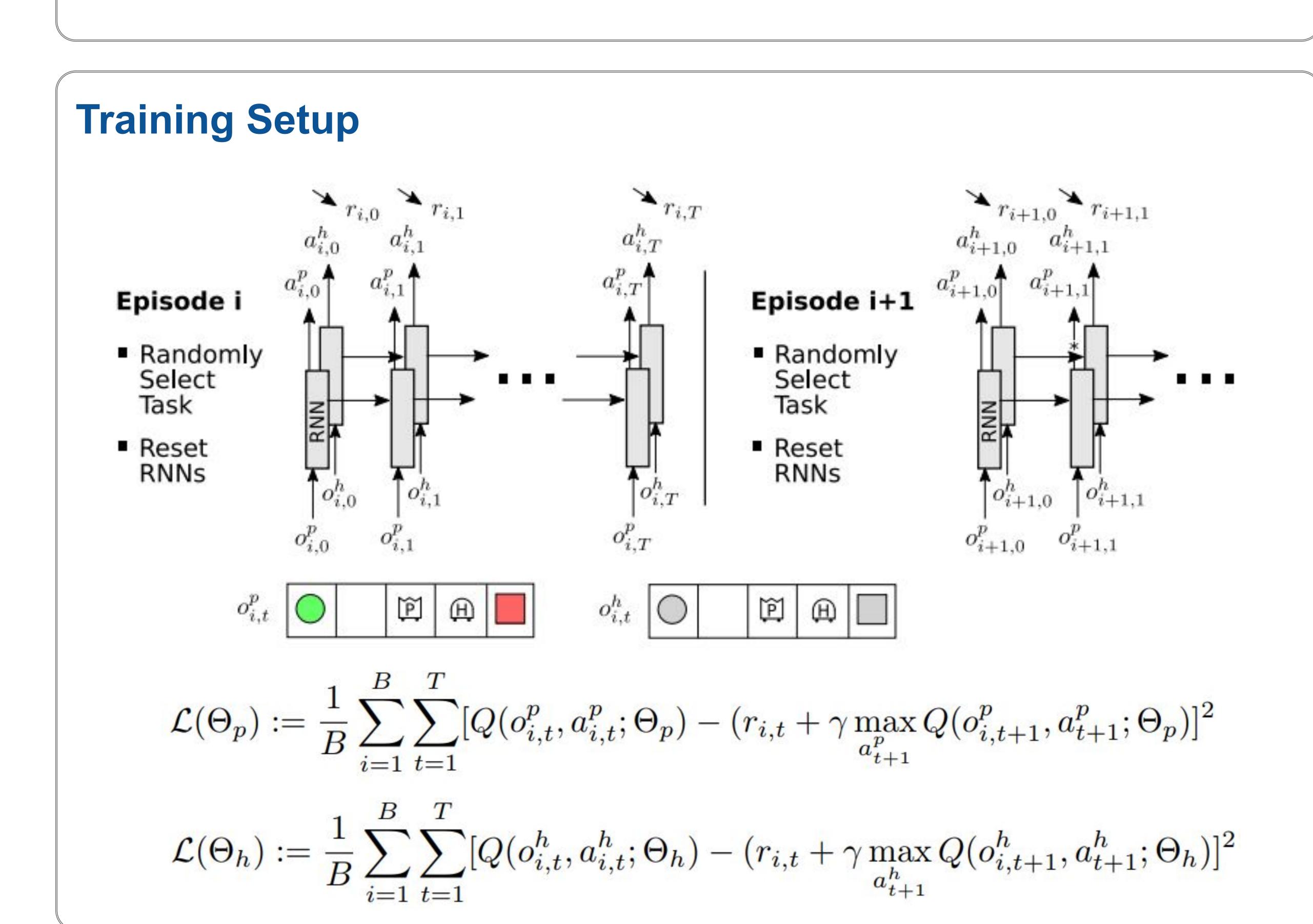


#### Motivation

- Learn a new task without a reward function or explicit demonstrations
- e.g. A "room straightening" robot interactively learning where to place items

#### Method

- Train in simulation on a distribution of tasks
- Train with a "prime" agent serving as a surrogate for the human
- Provide task specifics to the prime agent, but hide them from the "helper" agent
- Train the prime agent jointly with the helper agent



#### The Helper Helps

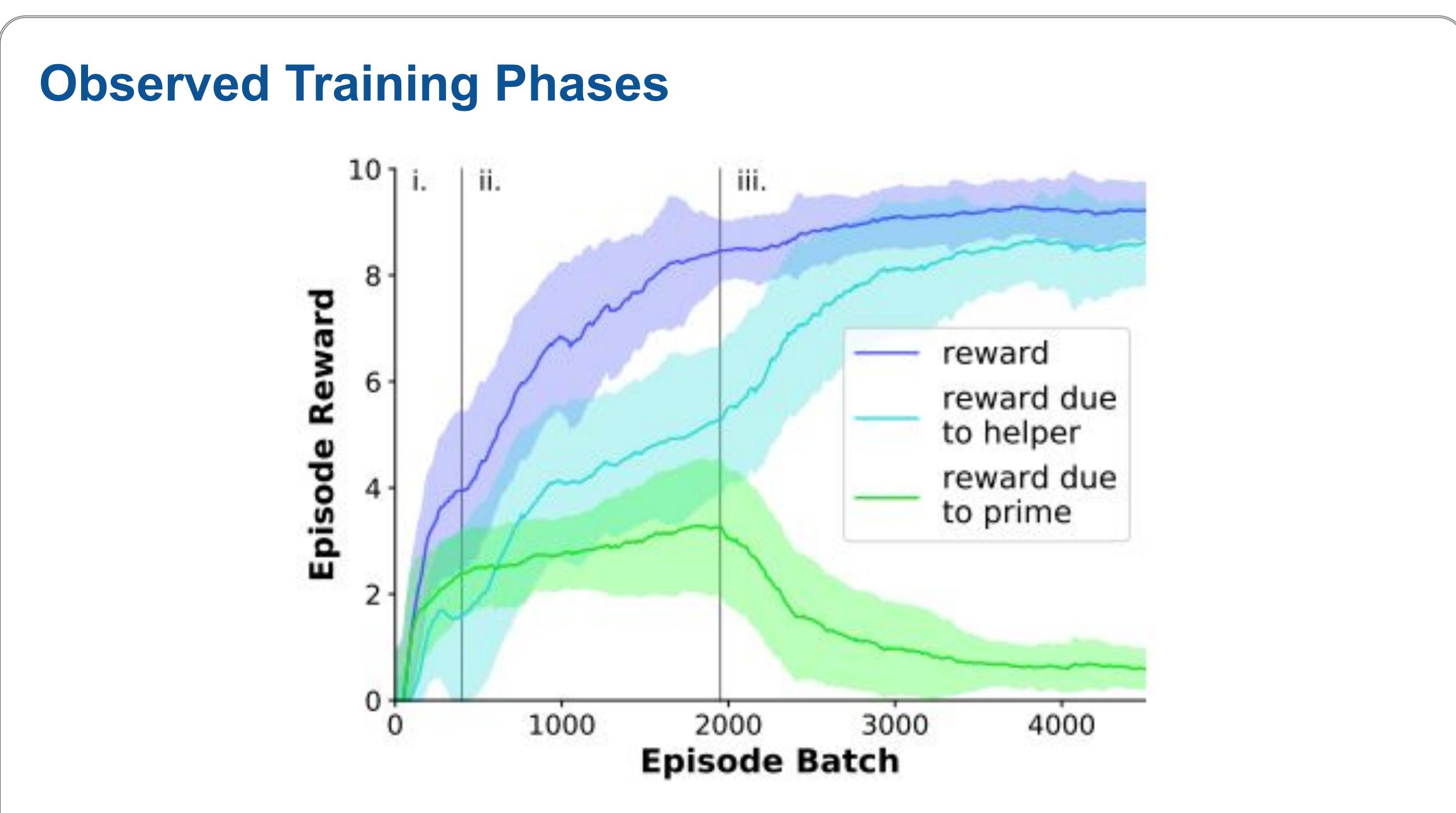
- Higher Episode Reward
- Fewer Prime Actions

	Without Helper	With Helper
Reward Per Episode	5.99	9.29
Prime Actions Per Episode	29.48	3.98

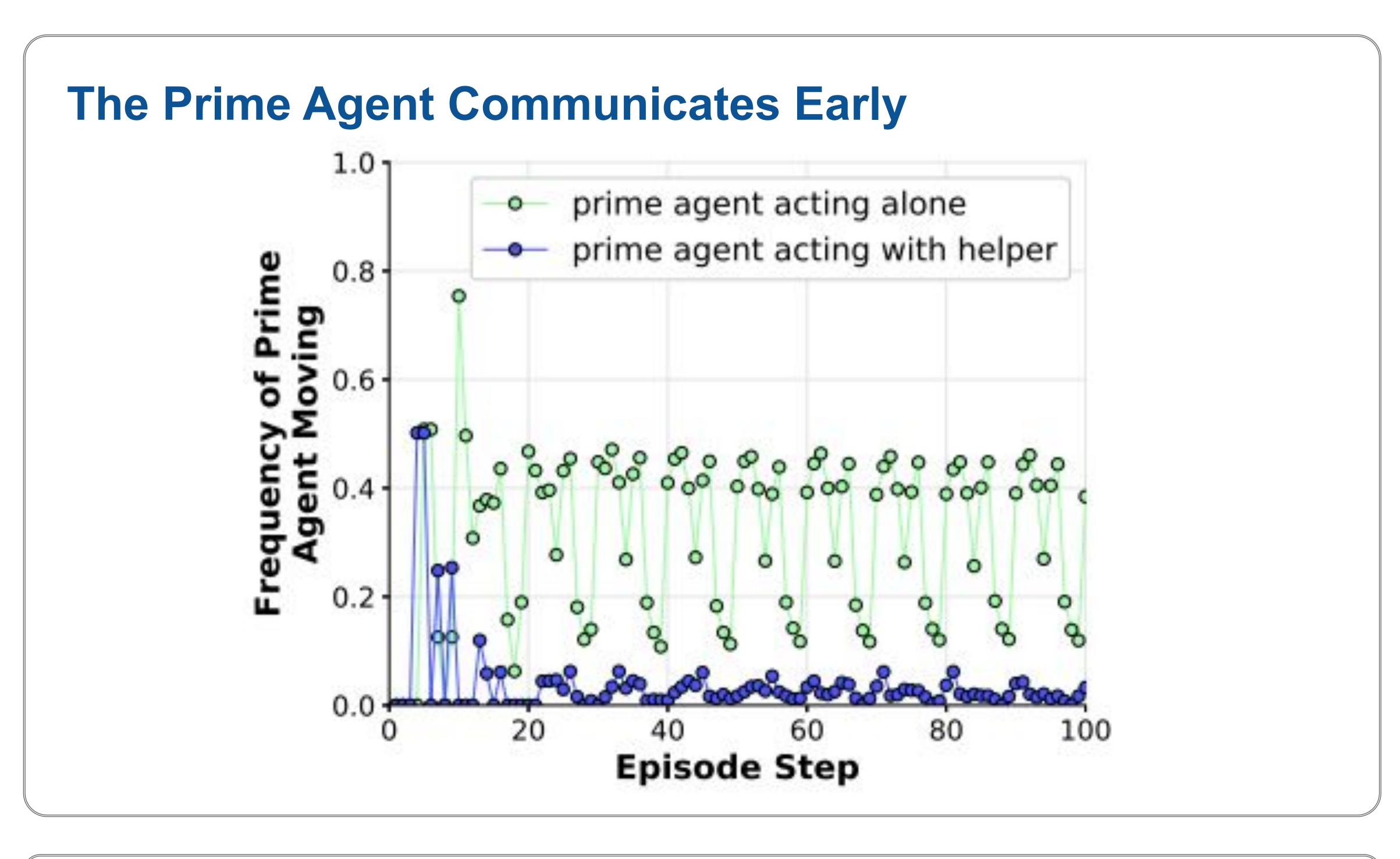
# \*Prime" and "Helper" agents foraging in a gridworld Observations binary vectors Prime agent observes "good" vs. "bad" \*\*Joint Rewards +1 "Good" object -1 "Bad" object -1 "Bad" object

-0.1 Prime motion

## **Emerged Communication** "Collect Circles" Įμ (a) An object from the "good" class appears first "Don't Collect Circles" (b) An object from the "bad" class appears first



- i. The prime agent learns to collect "good" objects
- ii. The agents learn to communicate and jointly collect objects
- iii. The agents learn to delegate to the helper agent



#### Contributions

- Proposed a method for training an agent to learn from interaction
- Demonstrated the method on a set of foraging tasks