

Reward changes are more disruptive than stimulus changes to implicit sequence learning of a community structure

Tejas Savalia¹ Andrew Cohen¹ Rosemary A. Cowell² David Huber²

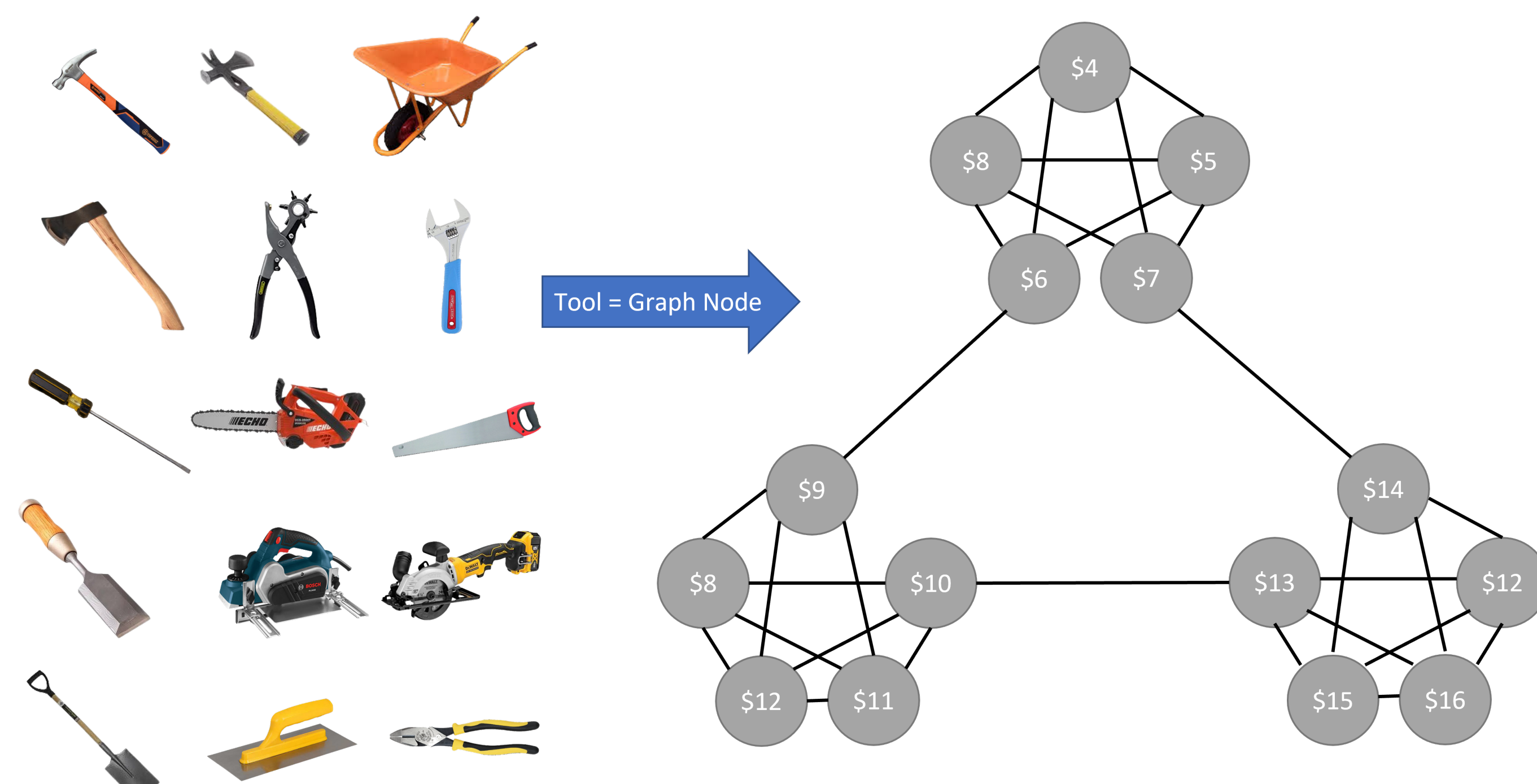
¹UMass Amherst ²CU Boulder

Introduction

People can implicitly learn a community-based transition structure in Markovian tasks. We consider whether the Successor Representation (SR) model of reinforcement learning can explain this kind of implicit statistical sequence learning (Mommenejad et al. (2017)).

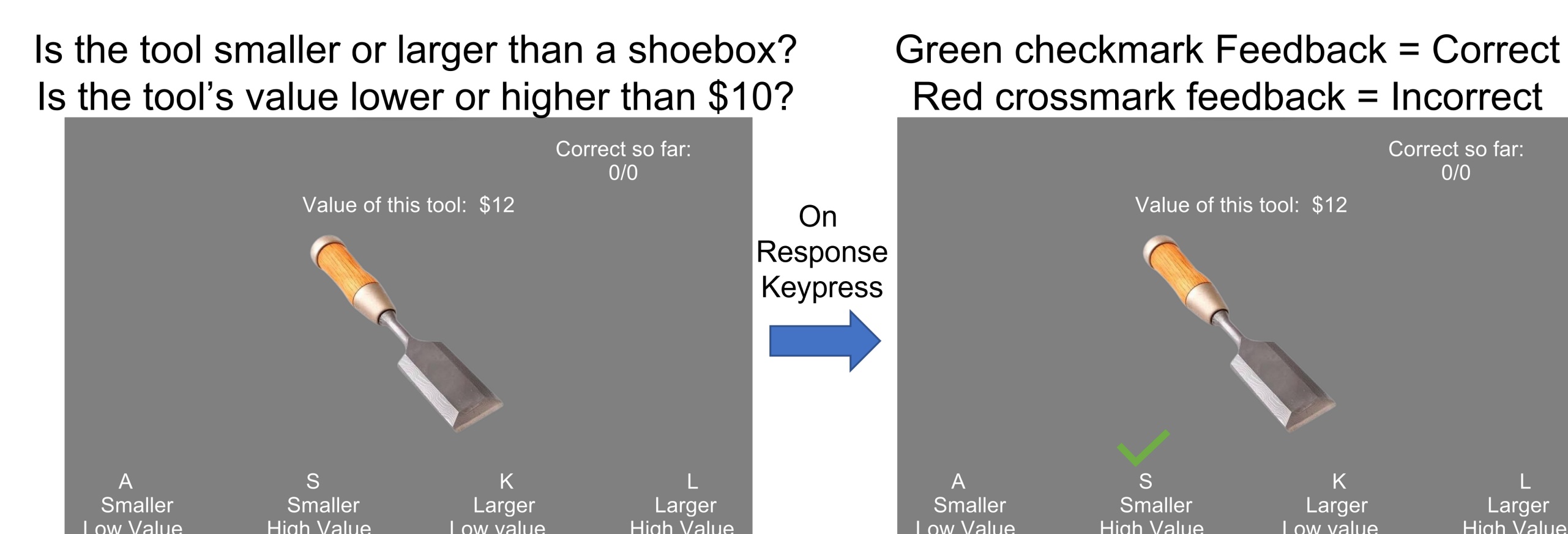
- The SR model learns transition structure through temporal difference learning, but learns rewards through single-shot, model-based inference.
- The SR model predicts that transition changes are more disruptive than reward changes.
- We tested this prediction in a task in which each state was defined by an image of a tool and each tool was assigned a dollar value.

Graph Structure



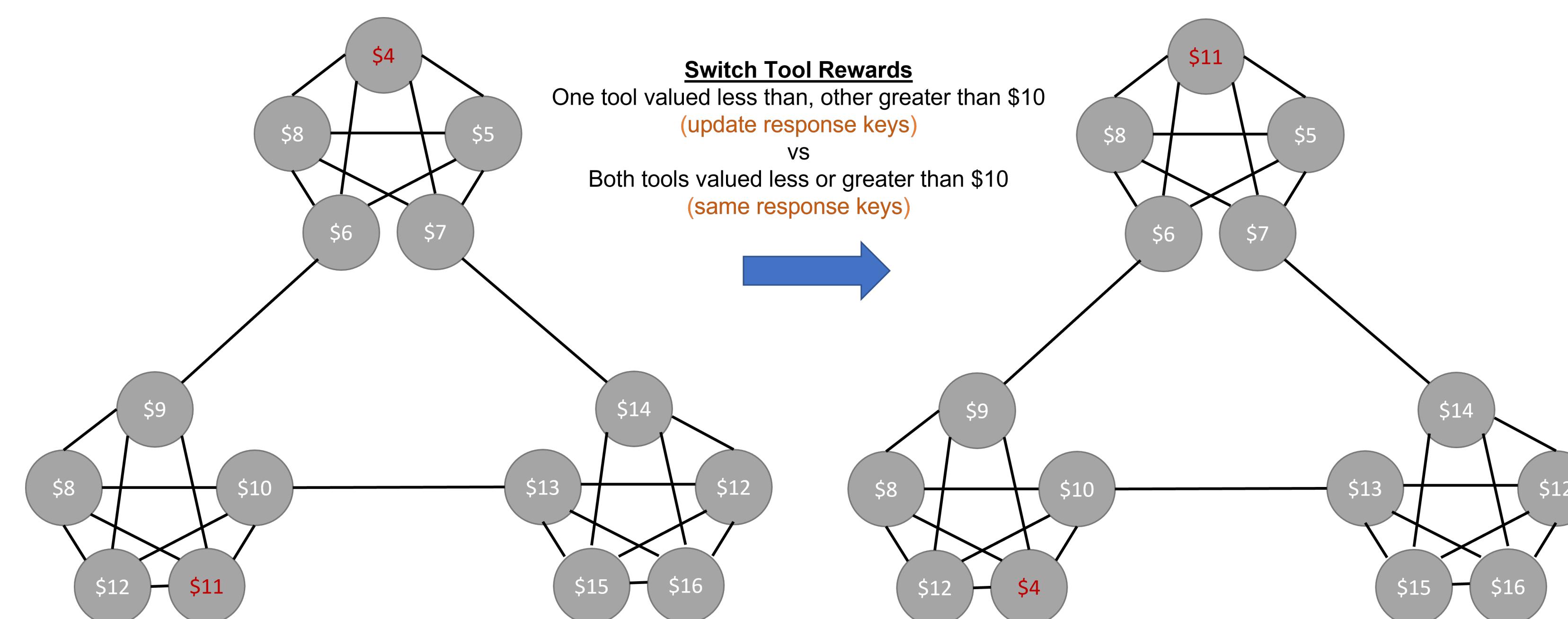
Tool arranged according to graph structure. (Schapiro et al. 2013).

Trial Structure



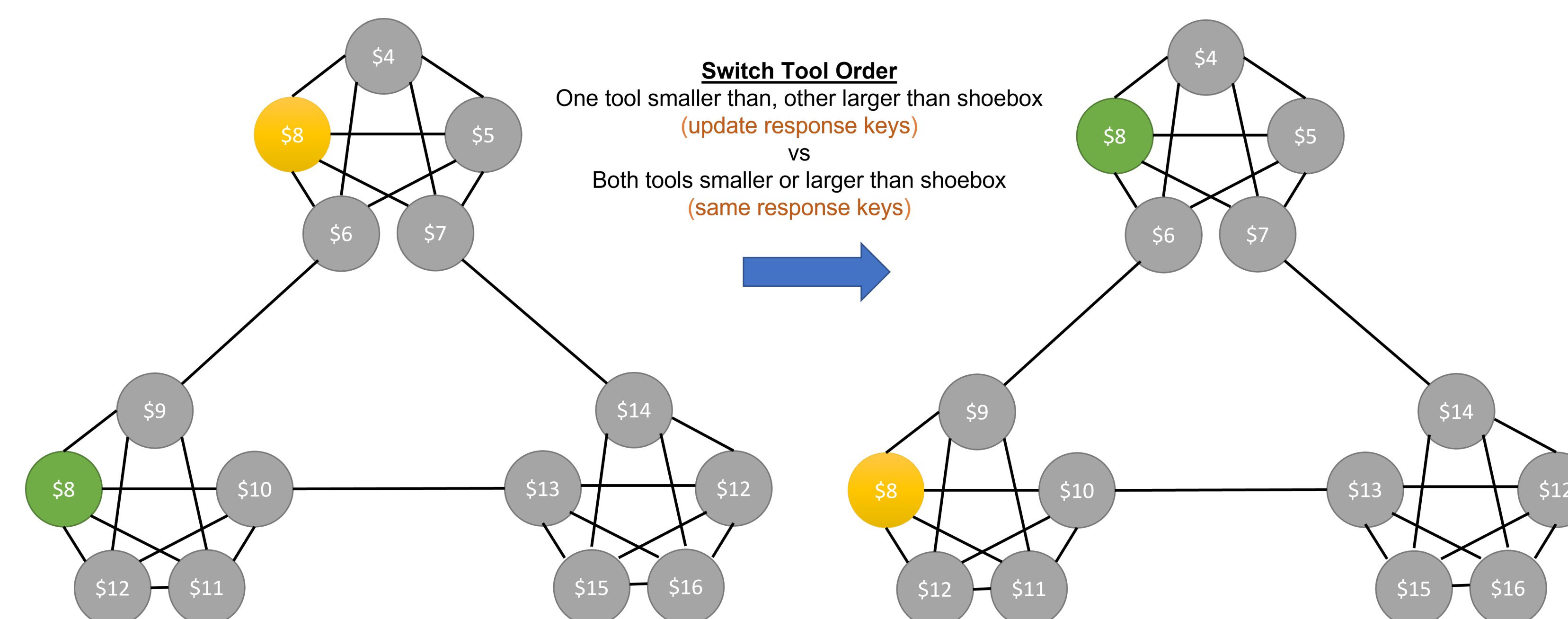
Stimulus order: Eulerian walk (each edge visited exactly once) through the Graph Structure.

Reward Switch



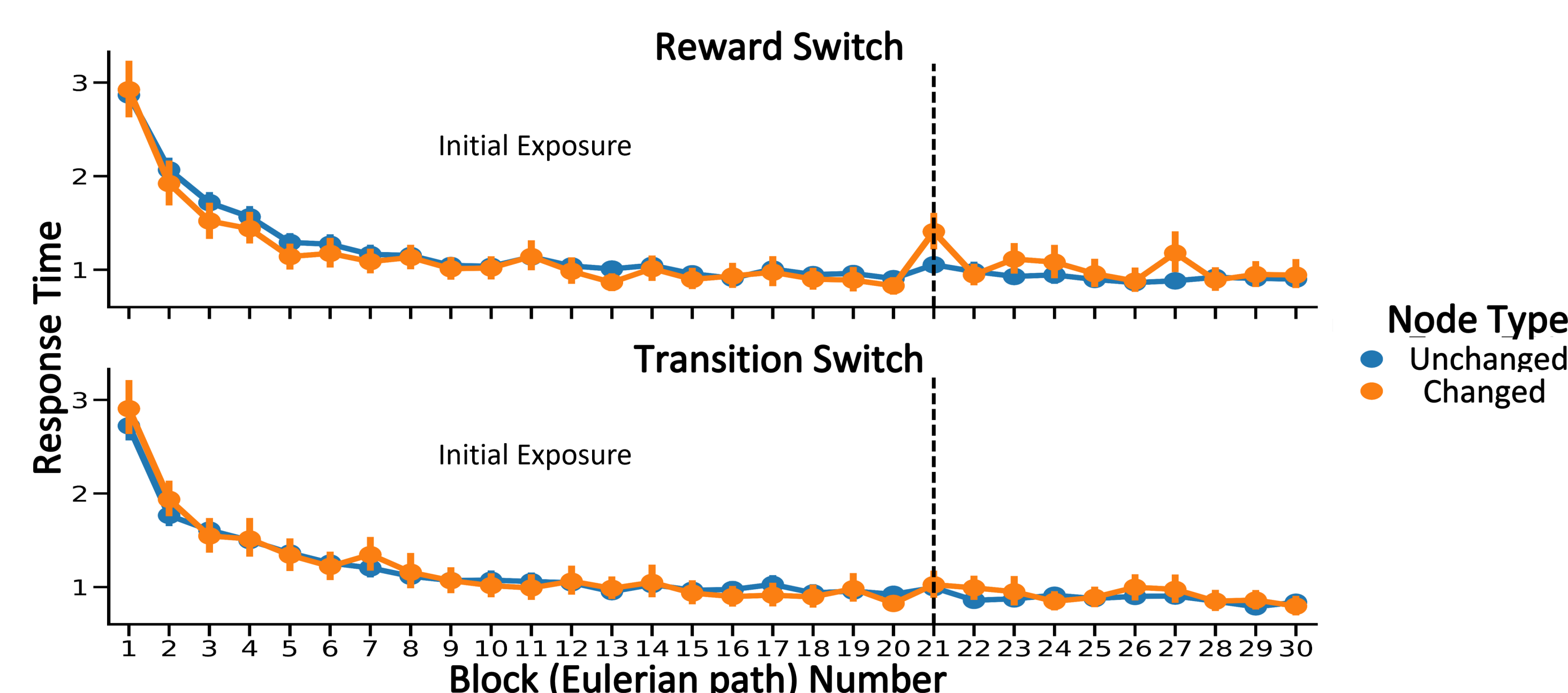
After 20 blocks of initial exposure, values of two tools were switched, keeping their state positions consistent. Relative values of tools was used to control for response key change.

Transition Switch

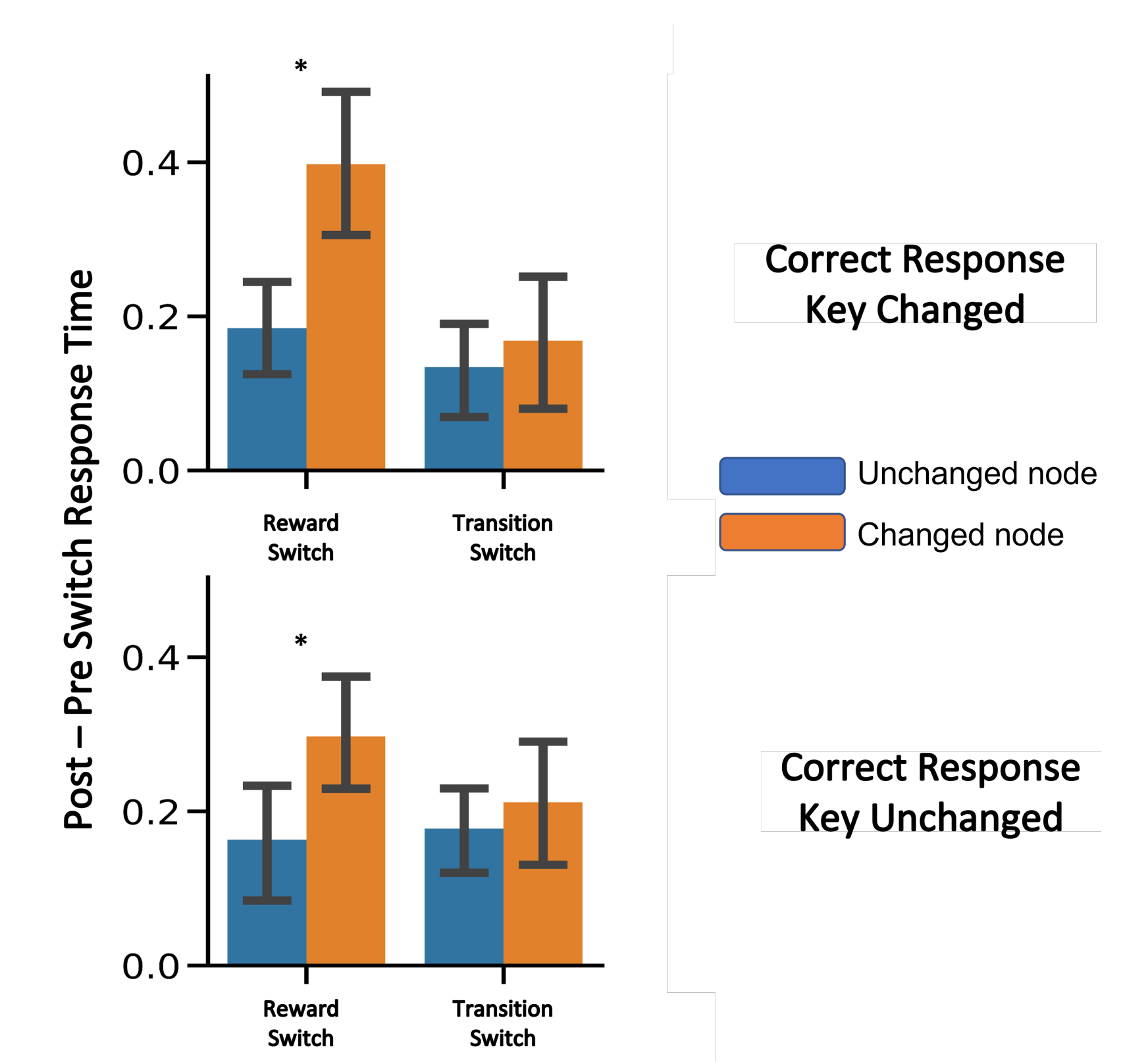


After 20 blocks of initial exposure, state positions of two tools of same value were switched. Size of tool was used to account for response key change.

Results: Response Times



Results: Change in Response Times



Results: Participants overall responded slower to the states that were changed. This slowness was more prominent for states with reward switch than with transition switch. The effect remains consistent regardless of whether the update required a new correct response key.

Key Takeaways and Future Directions

- In contrast to the SR Model predictions, reward change is more disruptive than transition change.
- Are tools 'states' or are values incorporated in a 'state' representation as features? We assumed that the state is solely defined by the displayed tool, but perhaps the onscreen values were also considered a defining feature of the state. If so, reward changes would act like transition changes. Additional work is needed to determine what defines a state.
- Does the SR model require multi-step decision making or forward planning tasks?

Acknowledgements

We would like to thank the cMAP-CEMNL and RDCL-DRM lab RAs (Aisling, Colleen, and Hannah) who helped collect data for this project.