

Motivation

I propose the use of *Model Human Learners* to aid designers in evaluating candidate learning interventions prior to conducting human studies.

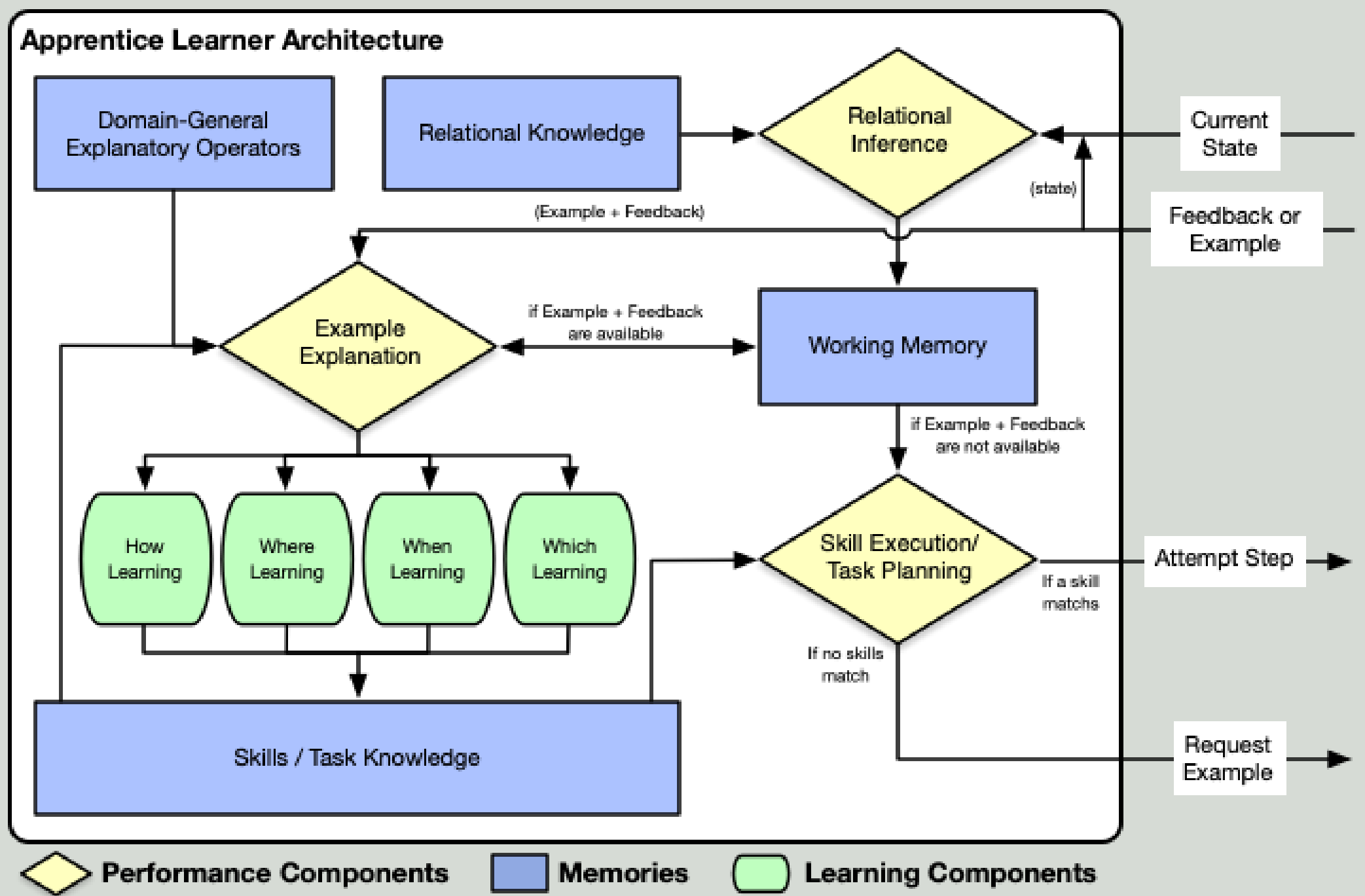
A Preliminary Model Human Learner

This work employs a computational model from the Apprentice Learner Architecture (MacLellan & Koedinger, 2022) that uses four learning mechanisms:

- **How Learning:** Searches for a sequence of explanatory operators to explain a provided example.
- **Where Learning:** Finds relational patterns to extract relevant information from the interface.
- **What Learning:** Learns the conditions under which learned skill should be executed.
- **Which Learning:** Learns a utility function to ranking matching skills for execution.

Unlike abstract learner models that fit functions to performance data, computational models of learning are *mechanistic*, simulating how knowledge is updated in response to practice and how performance changes.

To predict experimental results, the model is connected to learning environments and learns just like a human student, generating data that can be analyzed just like human data.



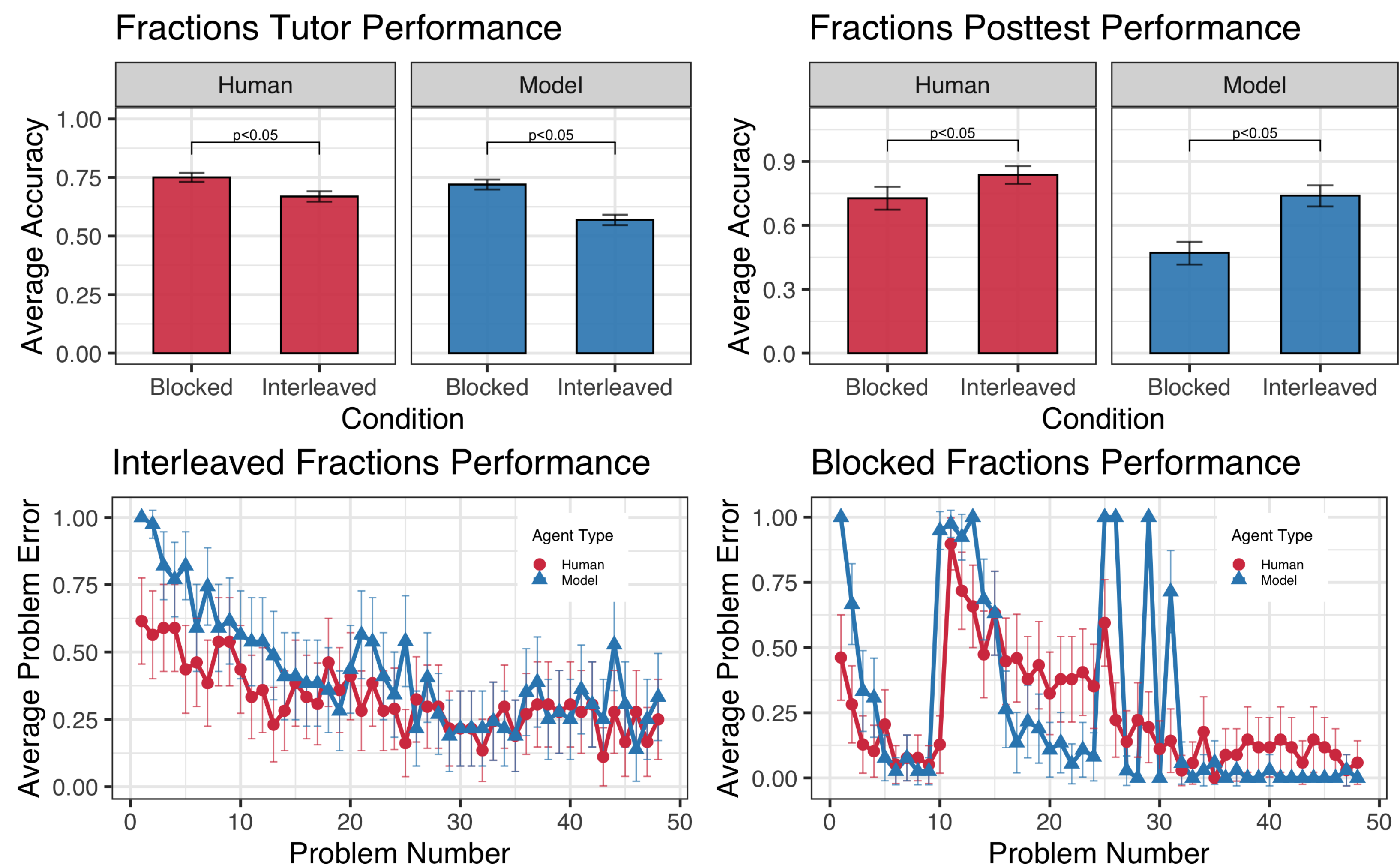
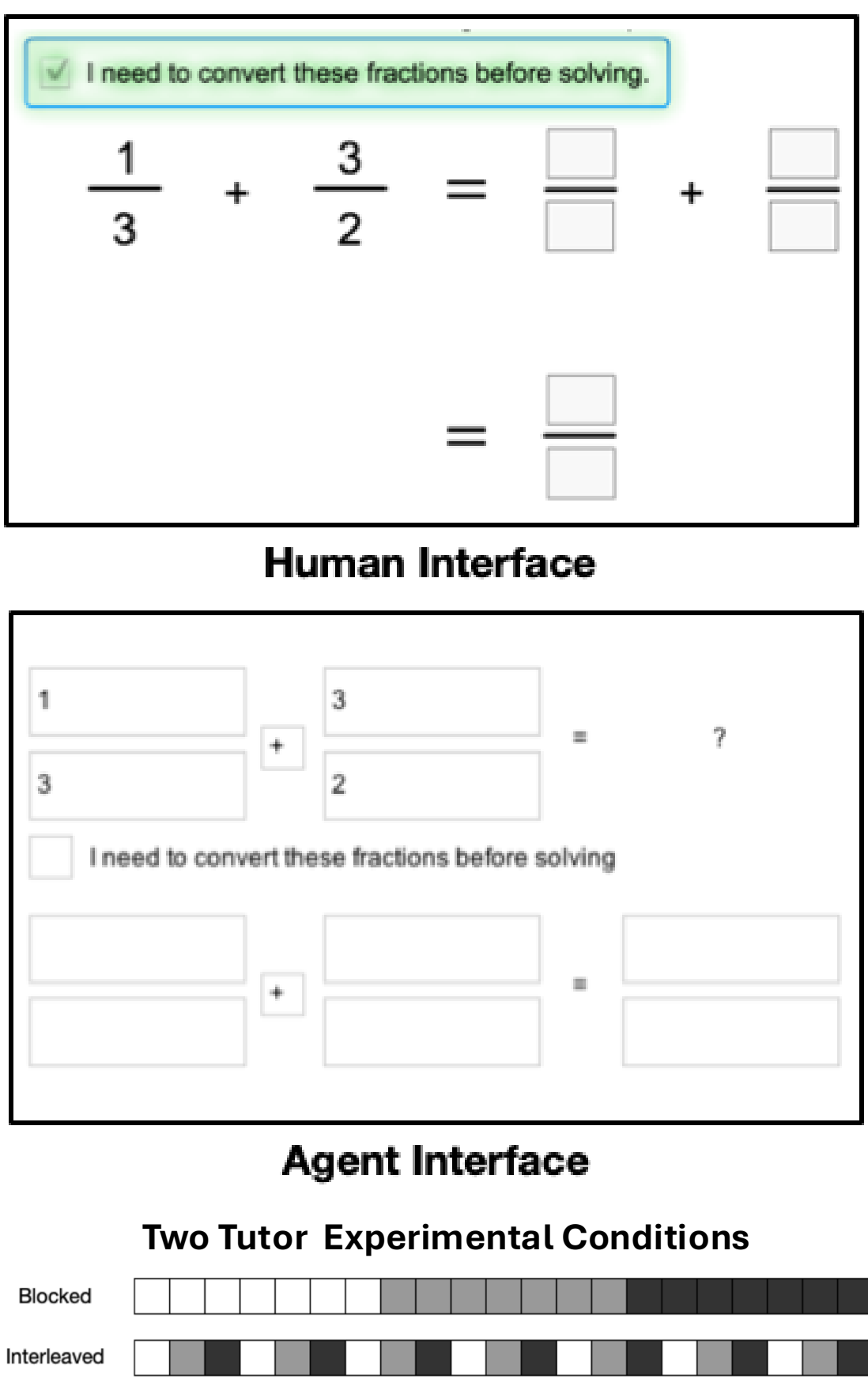
Fraction Arithmetic

Context: This study used the fractions tutor and data from Patel, Liu, and Koedinger (2016).

They showed that humans students have better tutor performance when instruction is blocked, but better posttest performance when it is interleaved.

Findings: Our model correctly predicts this main experimental effects and learning curve trends without ever seeing the human performance data.

Discussion: These results are a clear example of how tutor A/B experimental results can be predicted in a completely theory-driven way using a computational model of learning.



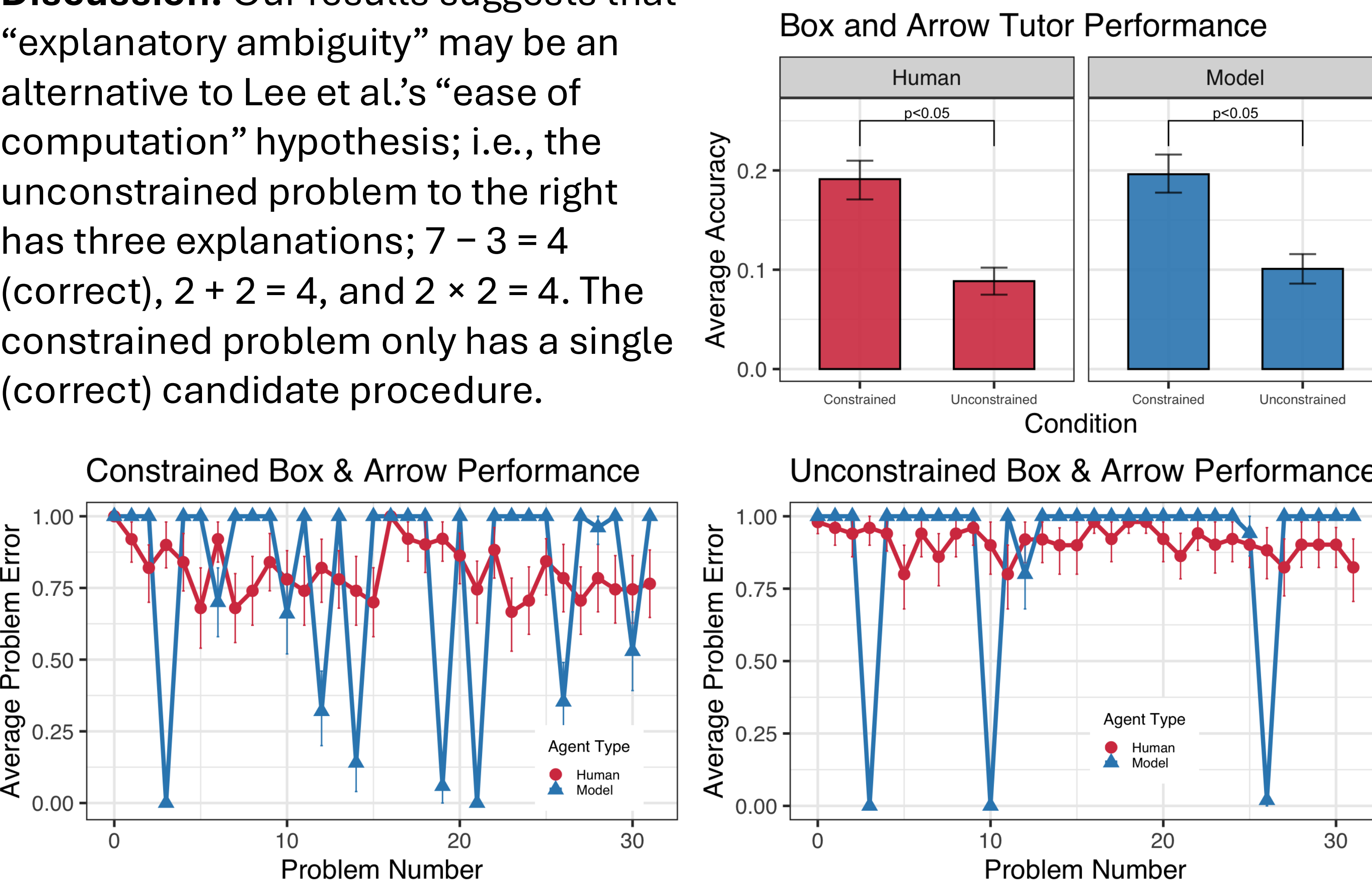
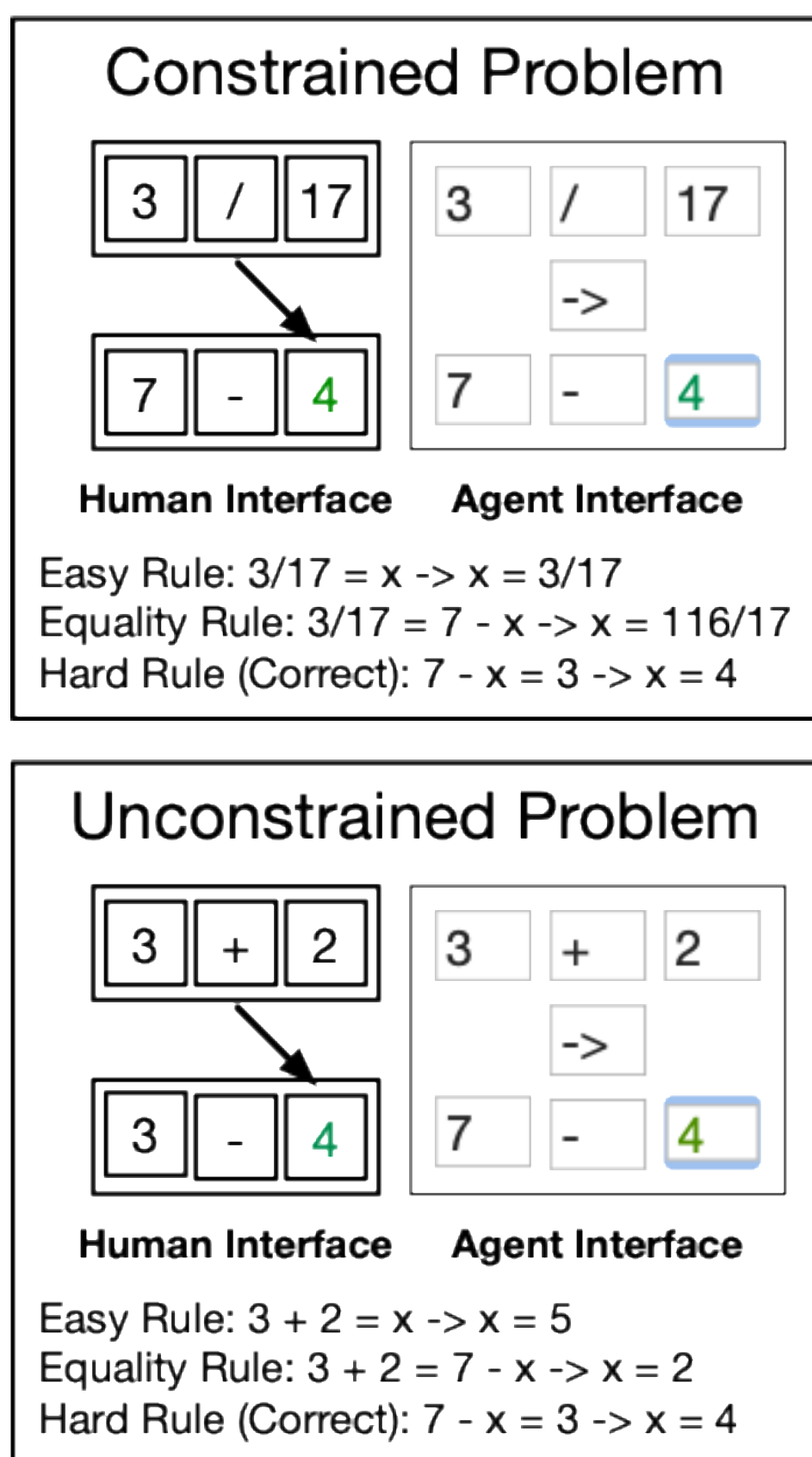
Boxes & Arrows

Context: This study used a boxes and arrows tutor and data from Lee, Betts, & Anderson (2015).

They found that students in the constrained condition had lower error rates than those in the unconstrained condition, hypothesizing that constrained problems bias students towards the correct procedure because they make the correct procedure *easier to compute* than the incorrect procedure (working with whole numbers is easier than fractional numbers).

Findings: Our model correctly predicts the main effect and learning curve trends, even though all the hypotheses are equally easy for the agent to use.

Discussion: Our results suggests that “explanatory ambiguity” may be an alternative to Lee et al.’s “ease of computation” hypothesis; i.e., the unconstrained problem to the right has three explanations; $7 - 3 = 4$ (correct), $2 + 2 = 4$, and $2 \times 2 = 4$. The constrained problem only has a single (correct) candidate procedure.



This work show that computational models of learning can:

- Accurately predict the main effects of two human experiments—one evaluating a problem sequencing intervention and the other testing an item design intervention.
- Generate learning curve predictions that closely match human trends *without access to human data*.
- Offer theoretical insights into why interventions work, challenging a prior hypothesis about learning from problem solving by Lee, Betts, and Anderson (2015) and suggesting a novel explanation of their results.

References

- Lee, H. S., Betts, S., & Anderson, J. R. (2015). Learning Problem-Solving Rules as Search Through a Hypothesis Space. *Cognitive Science*, 40(5), 1036–1079.
- Patel, R., Liu, R., & Koedinger, K. R. (2016). When to Block versus Interleave Practice? Evidence Against Teaching Fraction Addition before Fraction Multiplication. In *Proceedings of the 38th annual meeting of the Cognitive Science Society*.
- MacLellan, C. J., & Koedinger, K. R. (2022). Domain-general tutor authoring with apprentice learner models. *International Journal of Artificial Intelligence in Education*, 32(1), 76–117.



SCAN FOR PAPER