**Slider time**

Pros: ability not necessarily confounded with effort

Cons: people might expect to be interrupted (possible solution is to include option to pause); interpretation is a bit weird (women think they will want to practice more?)

**Time on task**

Pros: get an actual behavioral measure

Cons: ability might be confounded with effort – previous studies suggest gender difference in performance (possible solution is to include control for performance during task itself – although this might not be perfect); people might be interrupted (could give option to pause)

**Count for task**

Pros: ability not necessarily confounded with effort; get something closer to behavior than slider time measure

Cons: update budget to reflect poisson regression; they may not practice as many problems as they expect leading to issues of interpretation