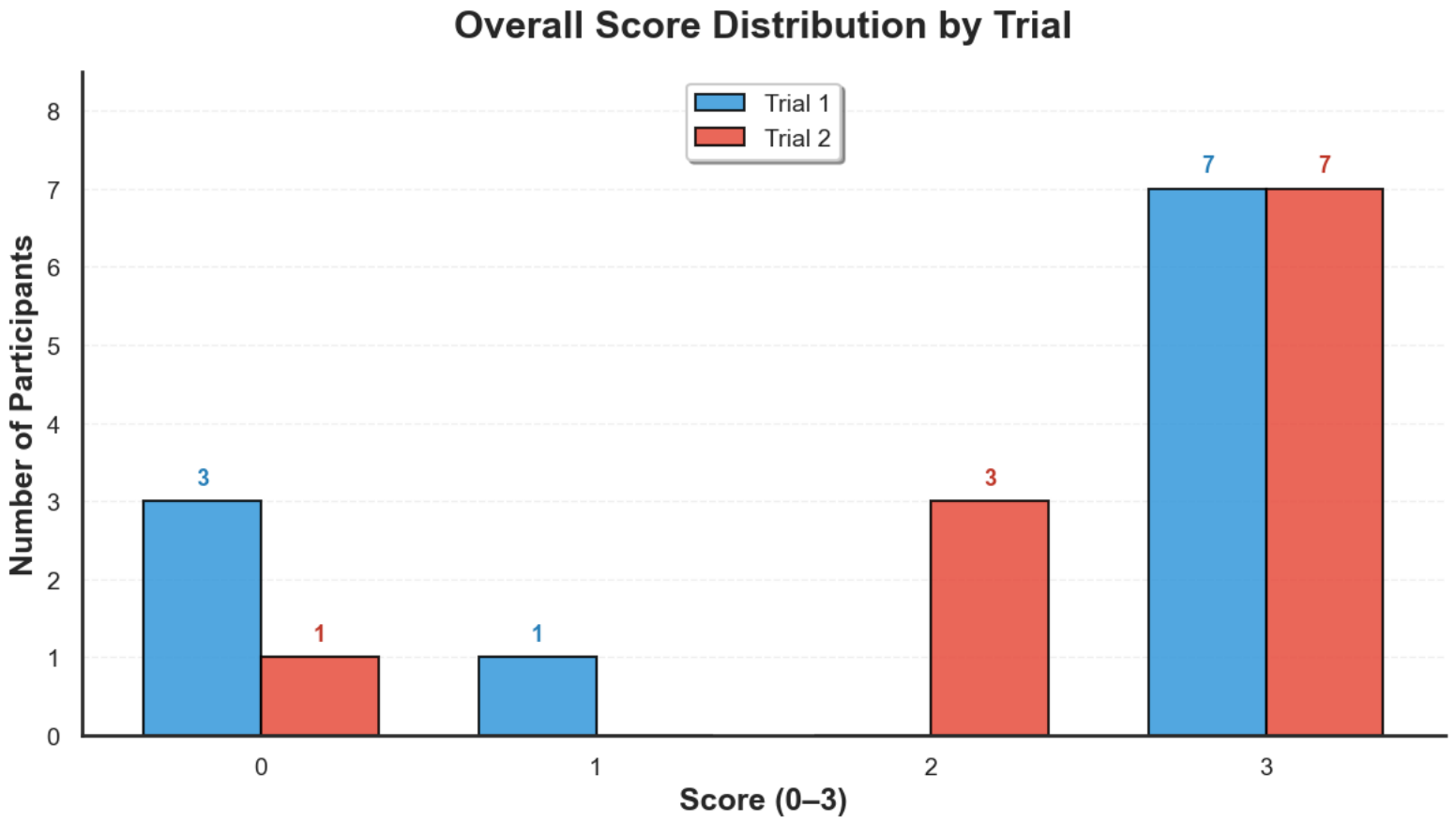
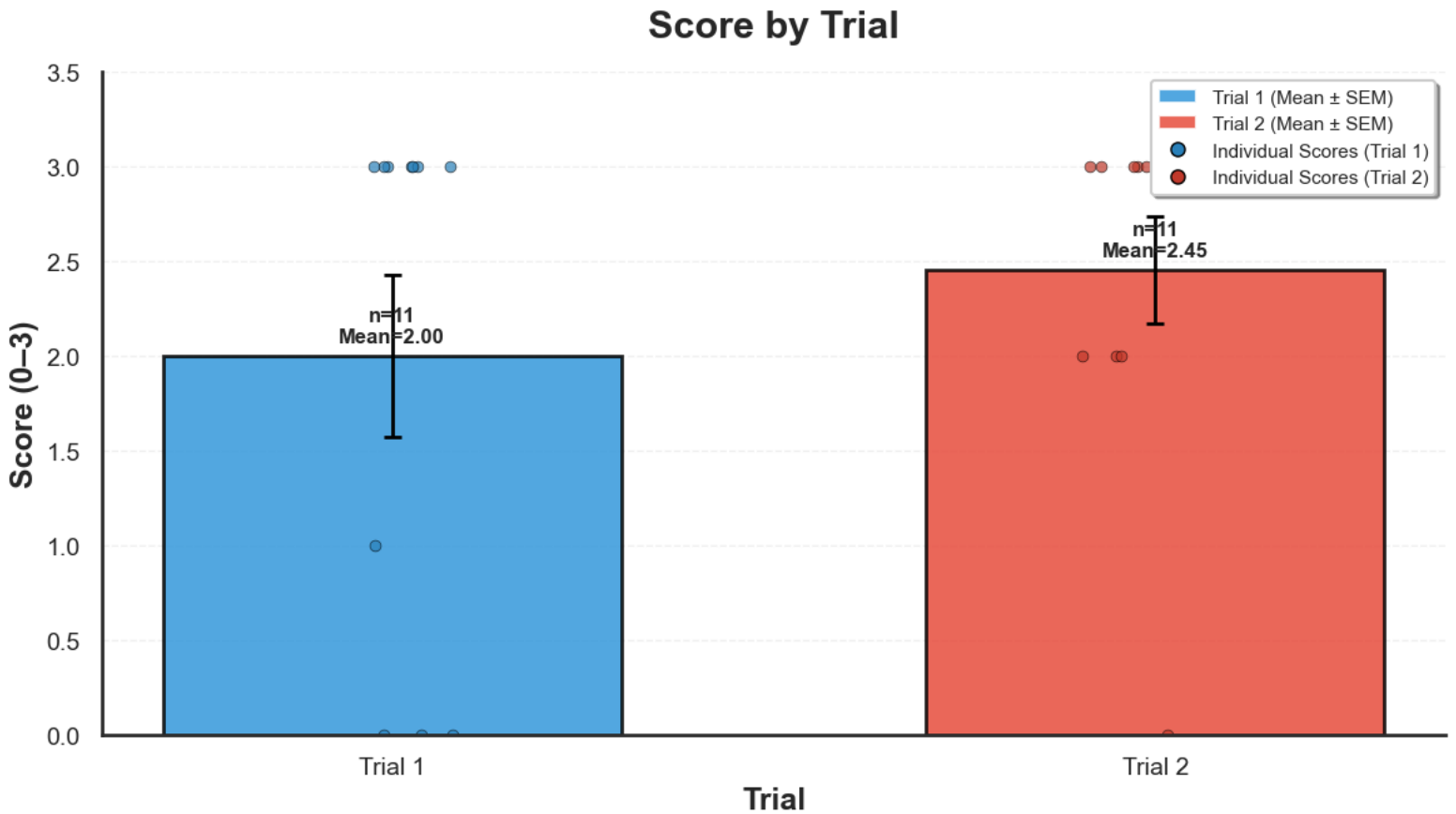
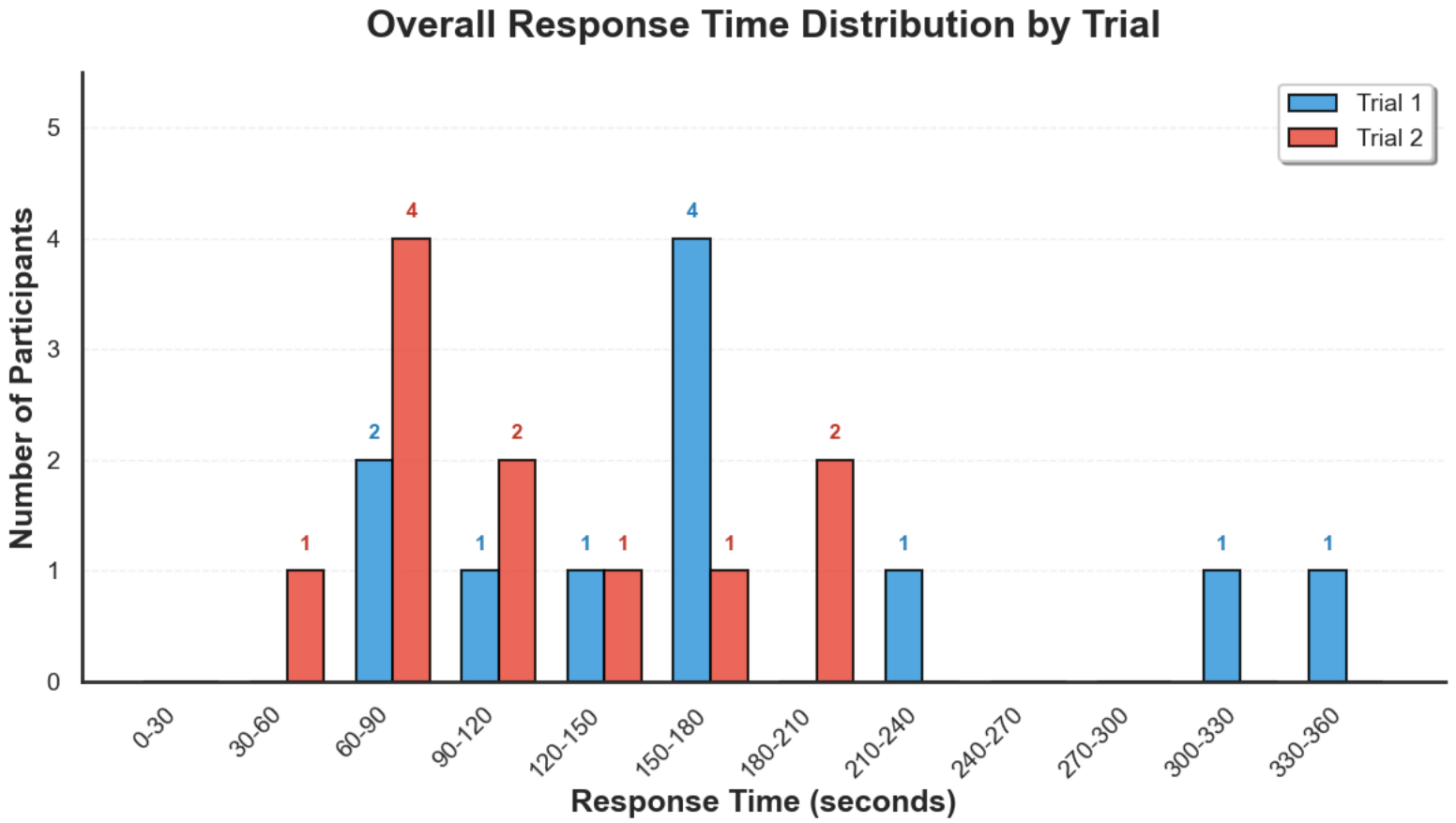
Pilot Data for Hanabi Task



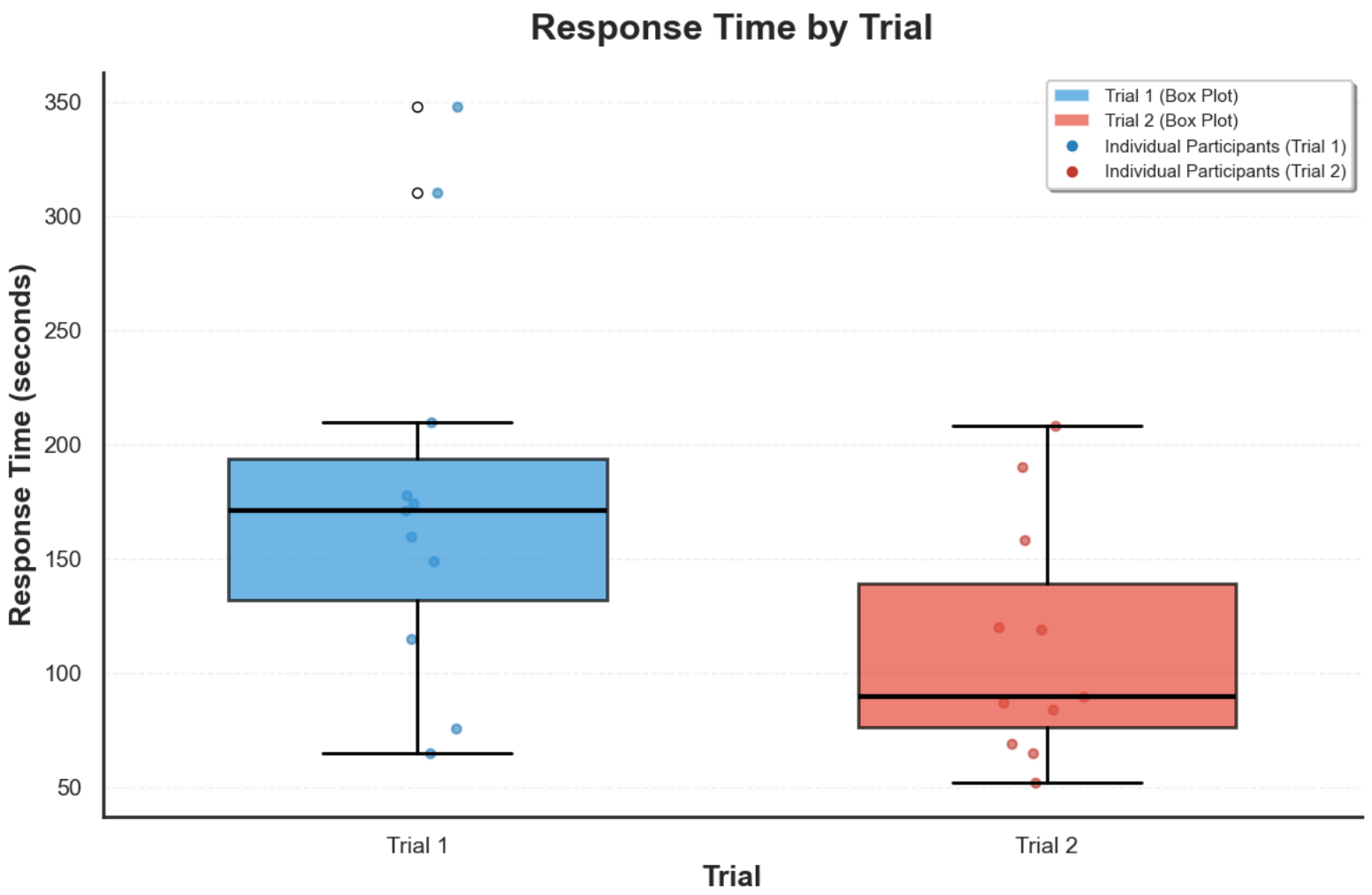
Trial 2 generally shows better performance.



Trial 2 generally shows better performance.



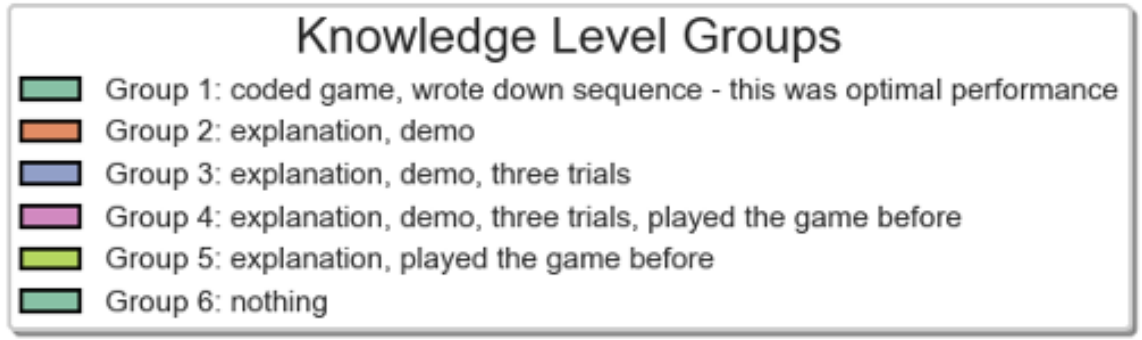
Trial 2 generally shows better performance.

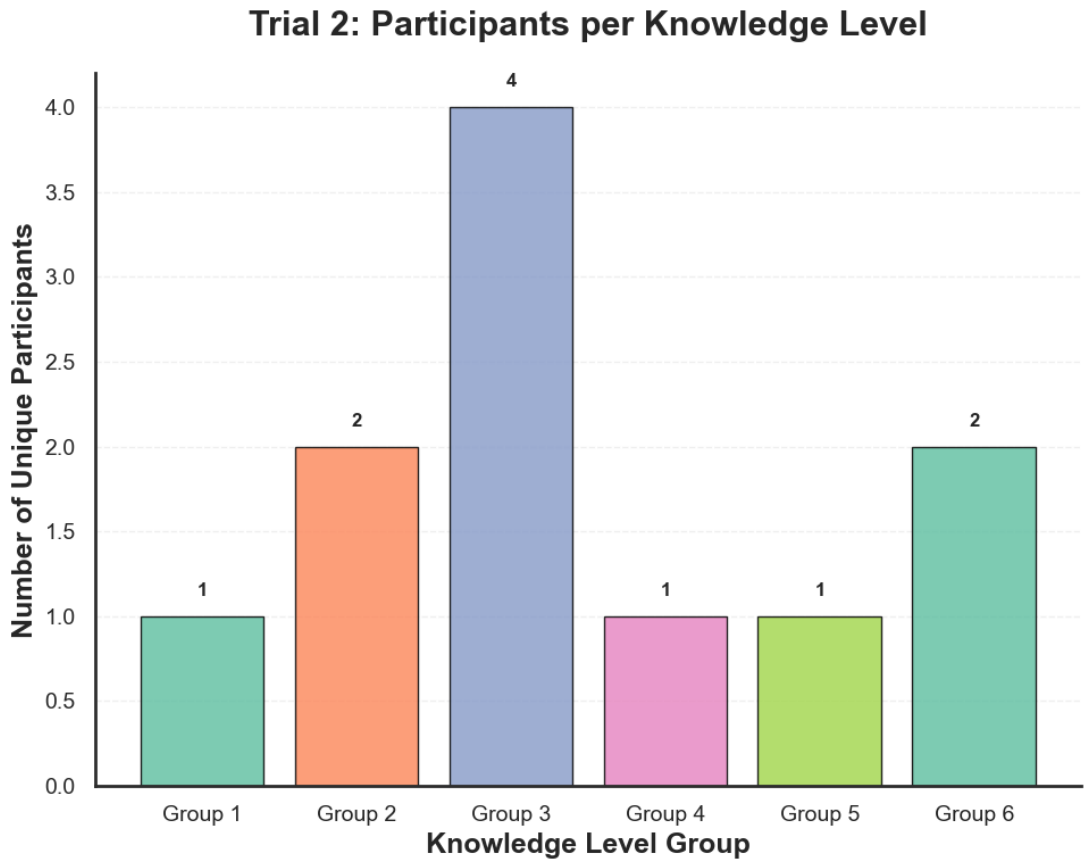


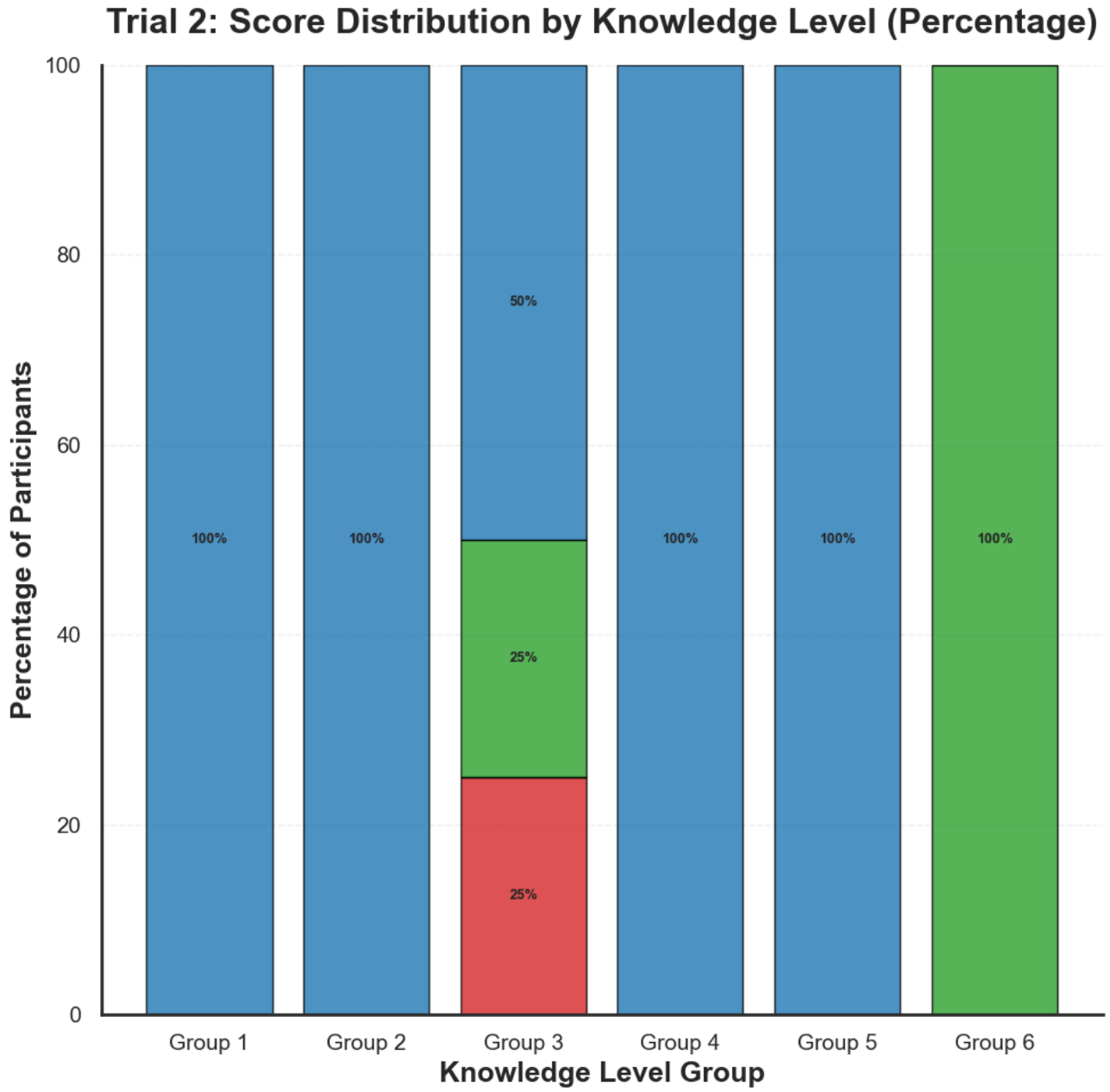
Trial 2 generally shows better performance.

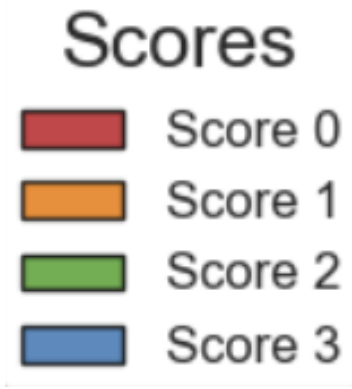
Knowledge level descriptors

- **coded game, wrote down sequence - this was optimal performance** - Kahini
- **explanation, demo** - had a lengthy explanation where I drew things out, saw a demo of me playing the game
- **explanation, demo, three trials** - had a lengthy explanation where I drew things out, tried the first round with me explaining it to them/demoing, then retried the task
- **nothing** - were just given the task and nothing else
- **explanation, played the game before** - the task was simply explained to them, but they had played the Hanabi game before
- **explanation, demo, three trials, played the game before** - had a lengthy explanation where I drew things out, tried the first round with me explaining it to them/demoing, then retried the task, had also played the hanabi game before





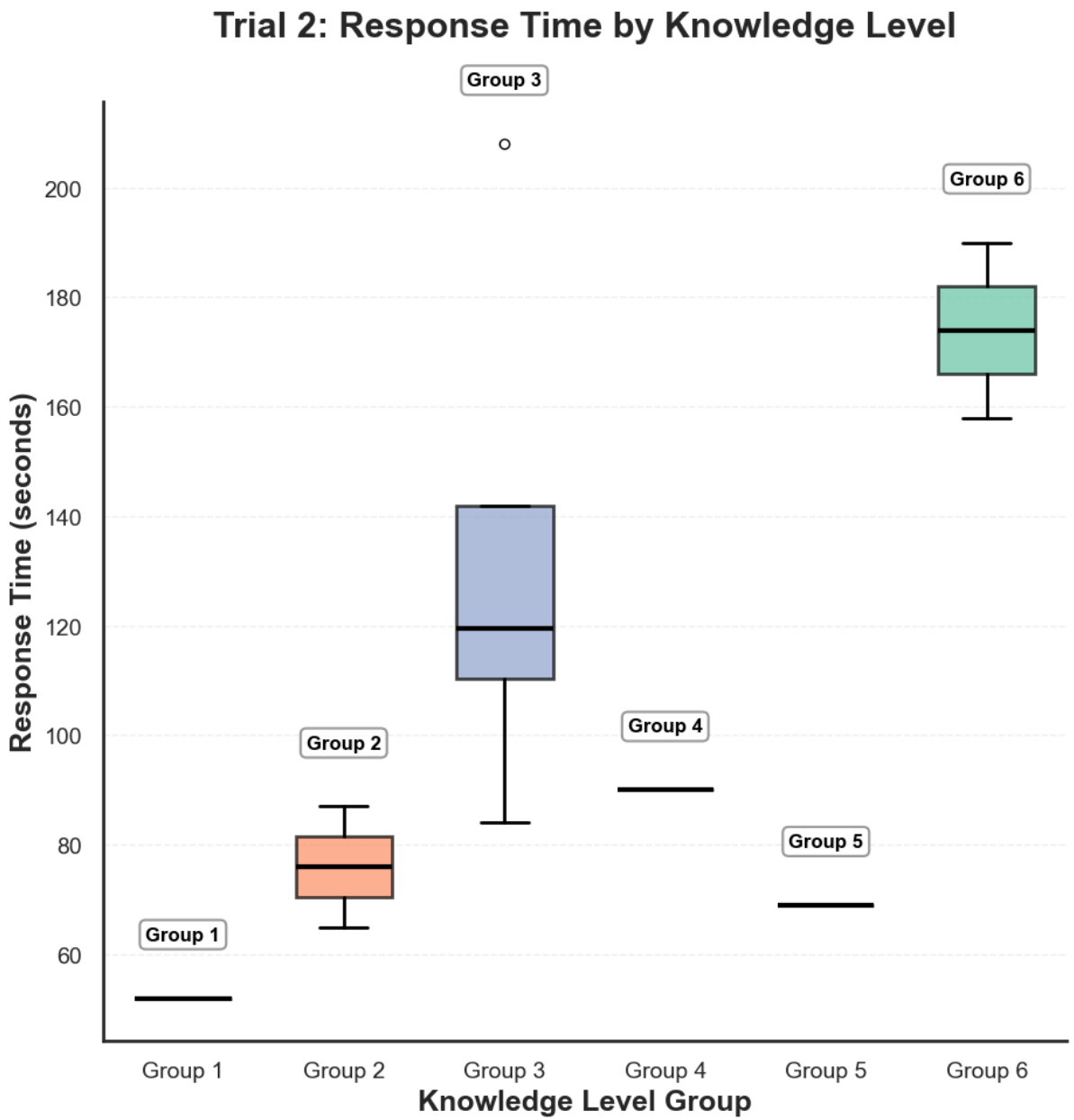


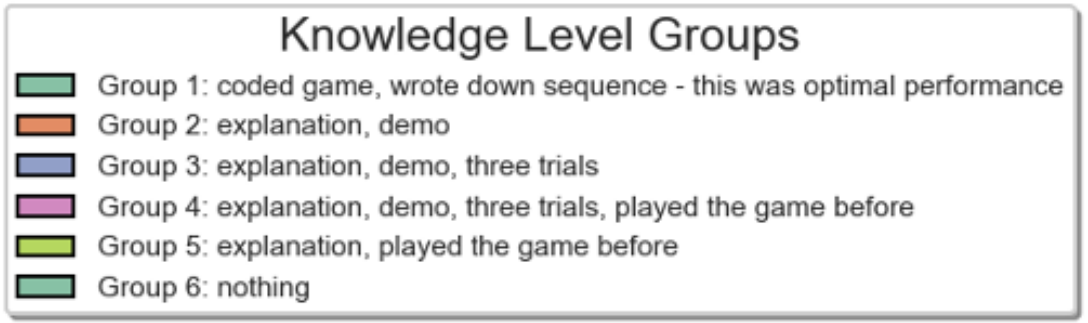


Knowledge Level Groups

Group 1: coded game, wrote down sequence - this was optimal performance
Group 2: explanation, demo
Group 3: explanation, demo, three trials
Group 4: explanation, demo, three trials, played the game before
Group 5: explanation, played the game before
Group 6: nothing

Group 3 had mixed results (took longer to understand the task), while group 2 (with no info) performed the worst.





The group with no information performed the worst. Those with explanation/demo or who had played before did best.

Common Complaints

- Cards disappeared once an action was selected.
- Too much to track at once.
- Was hard to remember the sequence.
- AI behavior wasn’t immediately clear -- when can I trust the AI/ is it optimal? Doesn’t always seem like it.
- This might be too complicated for cognitively impaired participants.
- Needed more guided practice.

Simplifications in this version

- Slowed hint timing for better clarity
- Added a demo round before the trials, made 2 trials
- Simplified AI logic to make it more predictable (eg: if it ever hinted on something, that card was in the sequence)
- Added a visual cue for player turn
- 1+ card from the sequence is always in play / if not immediately brought in on the next replacement

Suggestions for future versions?

- Include an interactive tutorial or guided first trial. Maybe a video?
- Better instructions during the task
- Make sure cards don’t disappear during actions
- Make AI logic more airtight/ clear to participants
- Maybe keep the sequence on the screen/ reduce the number of stimuli/ increase the number of practice trials?

General notes

- Since each task takes a minute (at the least), it might be complicated to get a lot of rounds it. Especially with ~6 minutes of explanation, ~5 minutes of practice, ~8 minutes of localizer – that gives us 25 minutes, so optimistically 25 trials.
- Everyone who performed the task has a graduate degree/ is working towards one. Might be more comfortable with a higher cognitive load, but still complained the task was confusing/ too hard.