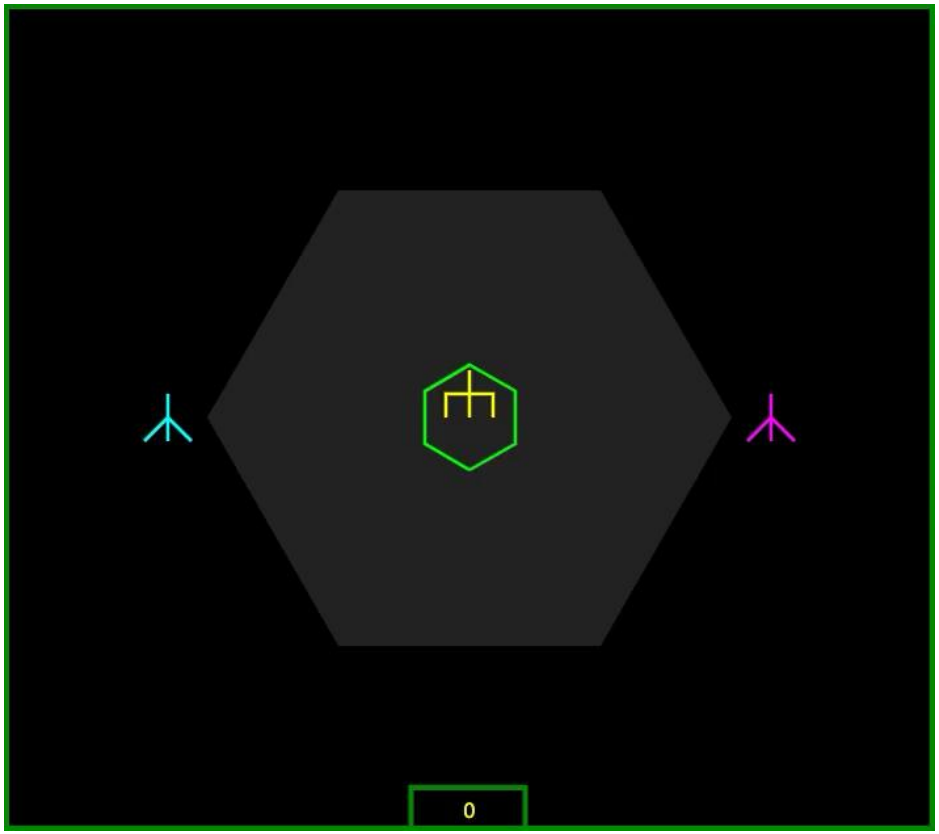
Model Tracing Open-Ended Tasks

Scale	Time units	System	World
10^7	Months		
10^6	Weeks		Social
10^5	Days		
10^4	Hours	Task	
10^3	10 Minutes	Task	Rational
10^2	1 Minute	Task	
10^1	10 Seconds	Unit Task	
10^0	1 Second	Operations	Cognitive
10^{-1}	100 Milliseconds	Deliberate act	
10^{-2}	10 Milliseconds	Neural circuit	
10^{-3}	1 Millisecond	Neuron	Biological
10^{-4}	100 Microsecond	Organelle	

Open-Ended Tasks:
Driving in traffic, problem-solving in uncontrolled settings, playing video games

Model Tracing:
Finding a sequence of model actions that matches a specific performance--Test of model, applications like tutoring.

Co-Op Space Fortress: An Open-Ended Video Game



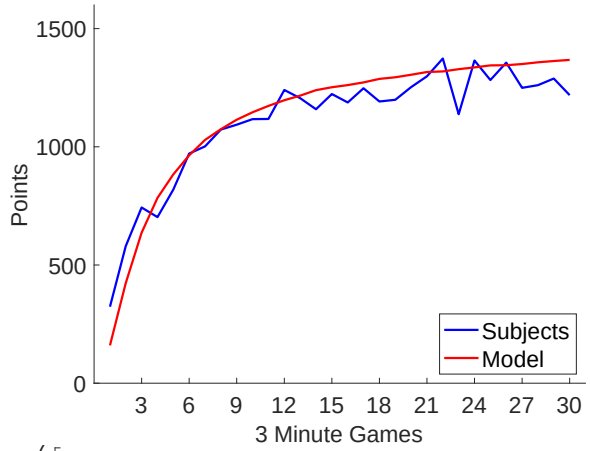
- Created by Cvetomir Dimov who also built an ACT-R model.
- Modified game to have fixed roles: Shooter and Bait.
- When the Bait flies into the hexagon, the Fortress tries to kill it.
- Firing at the Bait exposes its rear to the Shooter.
- The Shooter flies into the hexagon and can kill an exposed Fortress.
- If ships leave the hexagon the fortress will respawn.
- +100 points for a Kill, -100 for a Death, - 10 for a Miss.
- Major challenge is navigating in a frictionless space (keys for thrust, clockwise, counter-clockwise).
- No flight path ever repeats.

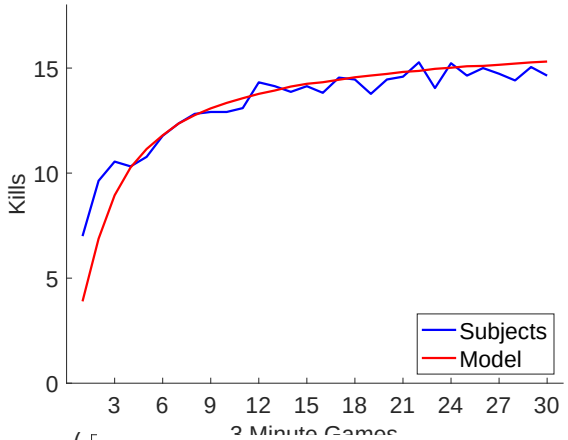
Dimov, C. M., Anderson, J. R., Betts, S. A., & Bothell, D. (in press). An Integrated Model of Collaborative Skill Acquisition: Anticipation, Control Tuning, and Role Adoption. *Cognitive Science.*

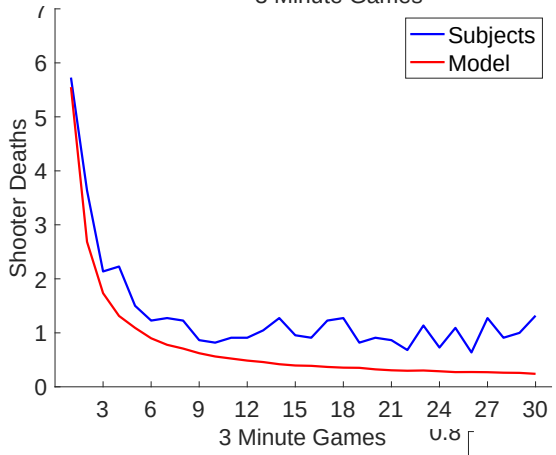
Key Features of Current Study

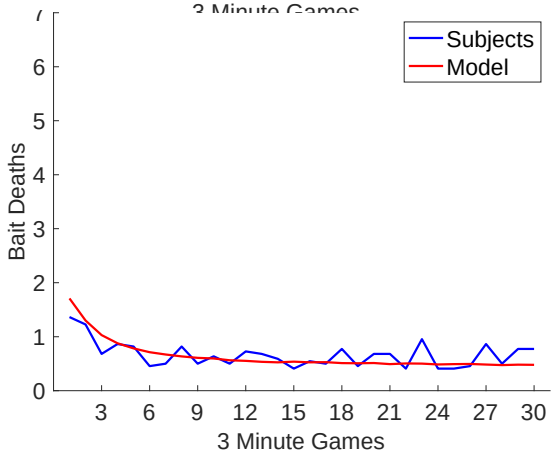
- The role of the bait was played by a bot which behaved like a highly skilled ACT-R model -- i.e. with skill, but showing human limitation and randomness.
- EEG was collected but will not be discussed.
- 22 “competent” subjects played 30 games.
- 7294 “competent” ACT-R models played 30 games, creating a library of 218,820 games.

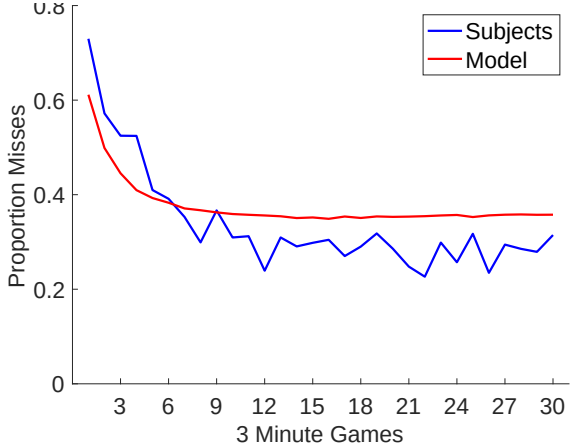
Subject-Model Comparisons





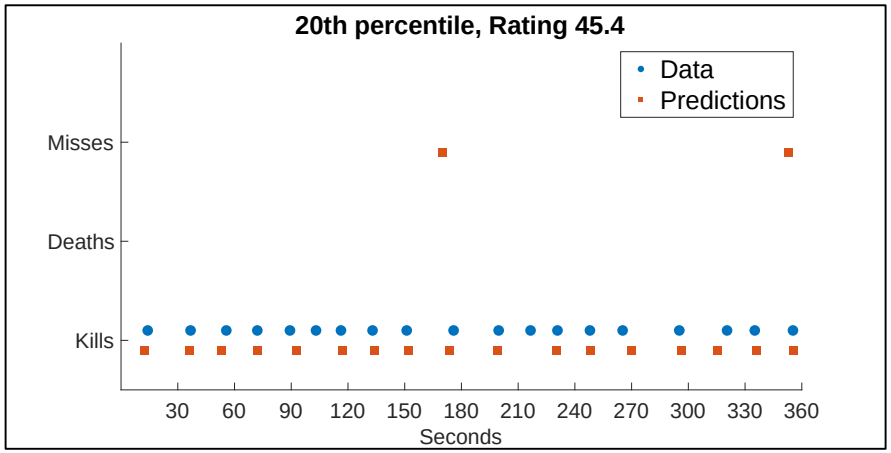


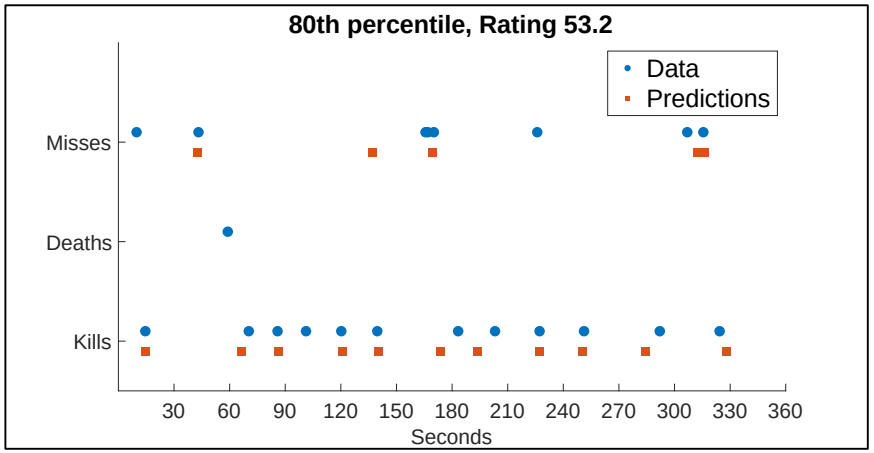




Model Tracing

- While there is correspondence at the level of average data, can the model produce a game that an actual subject played?
- Focus just on the kills, misses, and deaths: They correspond to the Unit Task Level
- No subject game or model game was like any other subject or model game. The closest match to subject games in the 200,000+ model games averages an error of 49



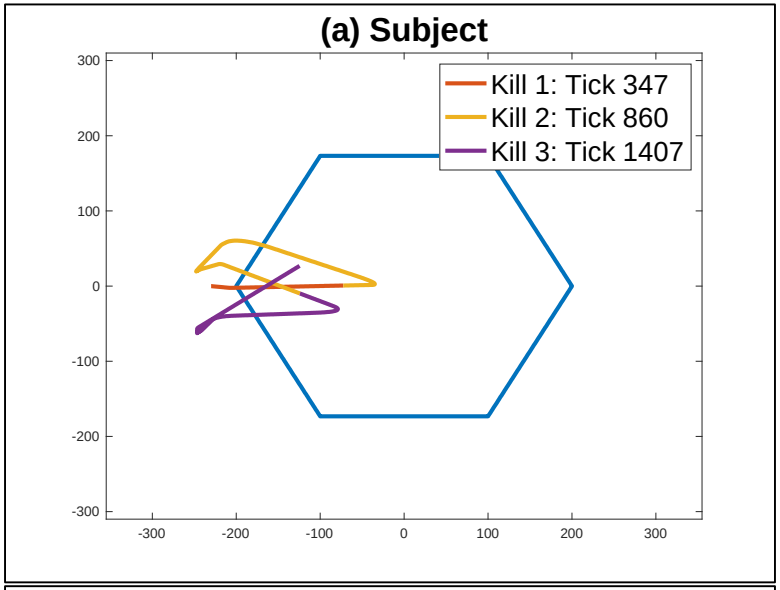


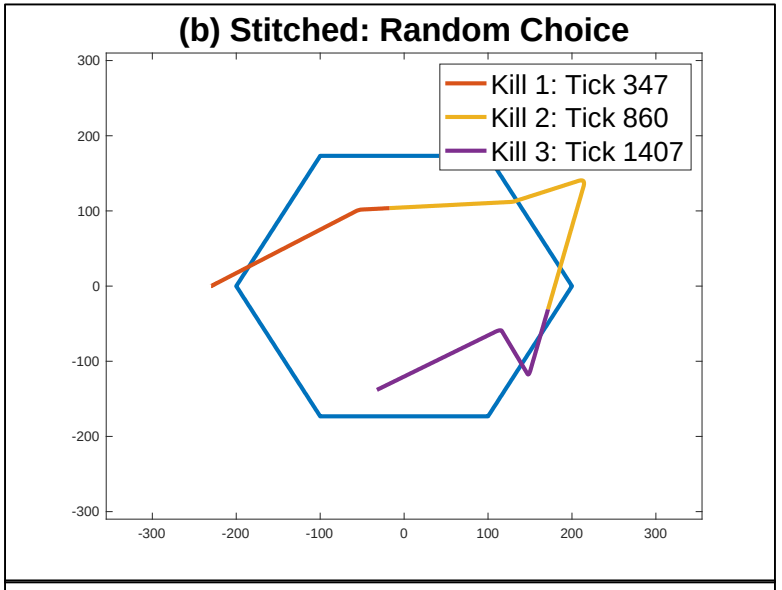
- Model Tracing 1: Repeatedly run model until it matches first event, freeze that point, run from there repeatedly until matches second point, etc.
- Problem: The model can be run to match first n events, be frozen in a state where it never can reach the n+1 st event.

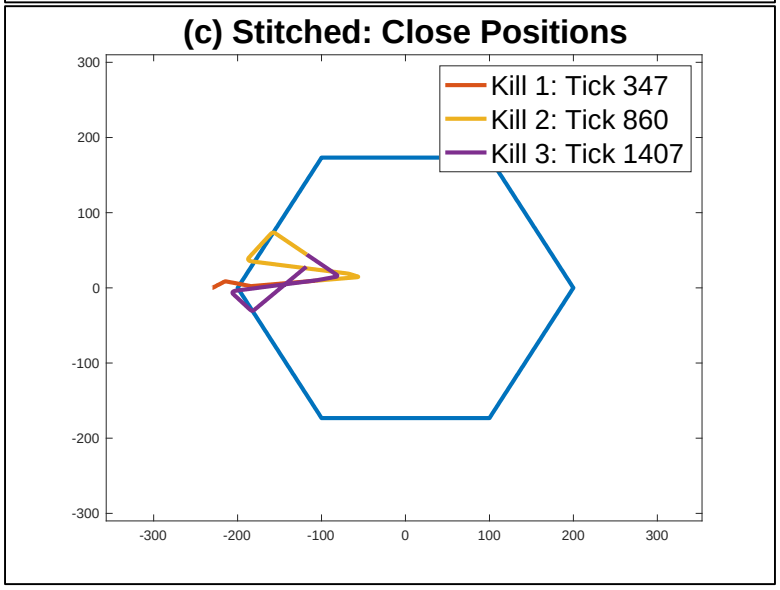
Model Tracing 2: Stitching

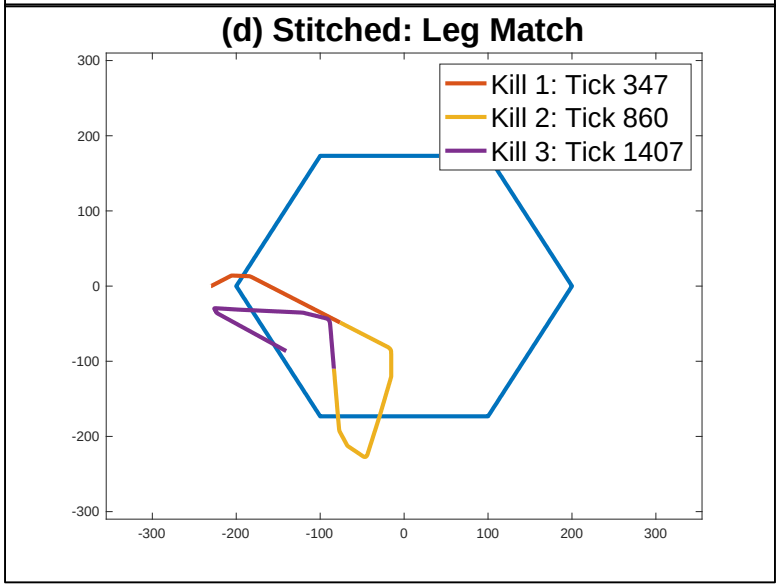
- The library of 200,000 games provides perfect matches to almost all the intervals between critical events in subject games.
- Just stitch **segments** from different games to produce a game that matches the critical events in subject games.
- **Segment** = the set of motor actions (turn-down, turn-up, thrust-down, thrust-up) that produced the transition from one critical point to another (and the reasoning behind those actions?).
- **Algorithm:** Given a game **fragment** to match the first n critical events, stitch in segments that match the length between the nth and n+1st events.
- **Problem:** the physical state (ship position, aim, trajectory) at the end of a segment from one game is almost never identical to the state at the beginning of a segment from another game.
- As a consequence, most attempts to stitch a segment into a fragment fail because the actions in that segment do not produce the next critical event.
- However, enough attempts succeed. The best fitting stitched game is off in its critical events by .03 game ticks (60^{th} of a second).
- Caveat: There is no guarantee, because of differences in physical state that the model would have actually chosen the same actions.

Example Flight Paths



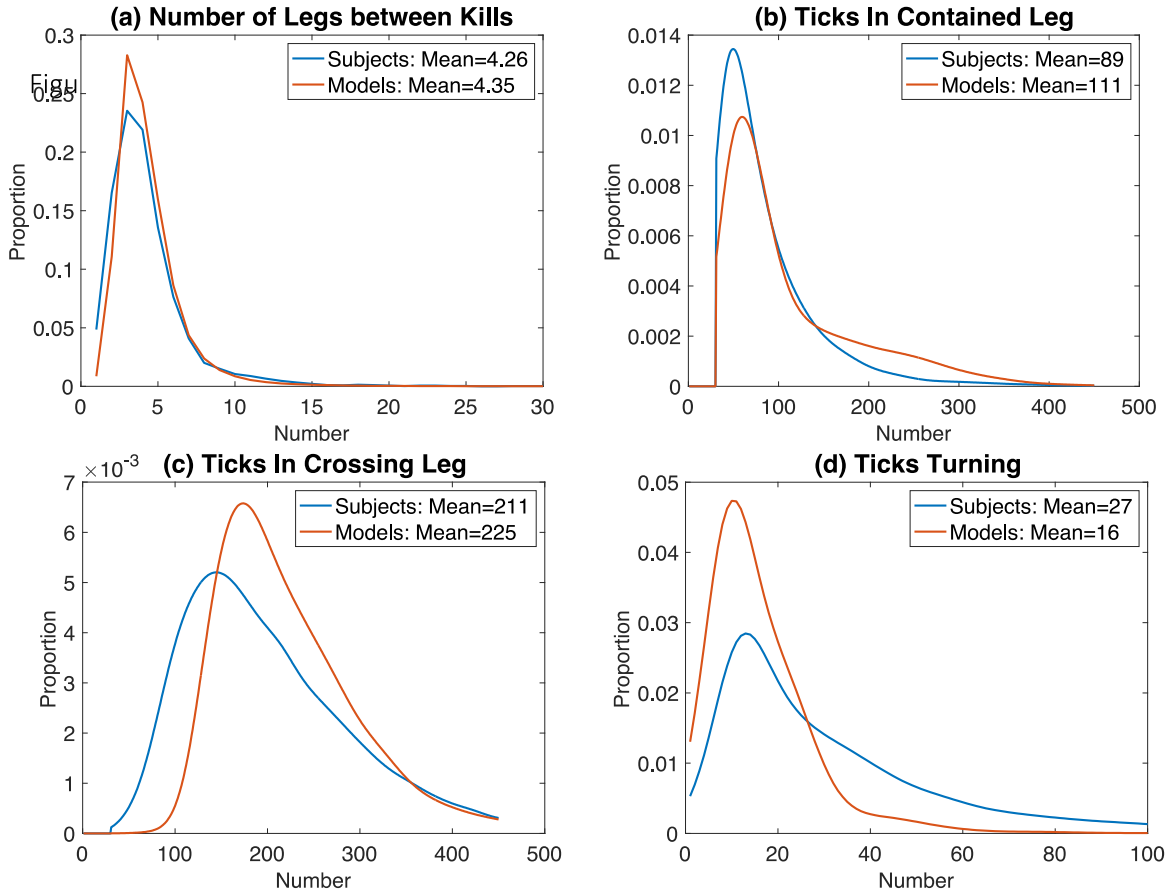






Legs of Flight

- Legs of flight correspond to Operations (below Unit Tasks in Cognitive Band) -- steps to achieving a goal such as leaving hexagon, going back in, flying to open back of fortress, correcting to flight paths, etc.



- Substantial overlap, but model has few short crossing segments or long turns.

Matching Legs of Flight

- Leg Markers: Ticks on which legs begin or end.
- Match: A marker in a stitched flight is considered to match a marker in a actual game if it is within 30 ticks.
- Choose stitchings with the highest match.

Selected Stitched Games: d-prime = 1.37

		Marker Present	Maker Absent
Actual Games	Marker Present	47,563	17,891
	Marker Absent	20,291	71,935

Random Stitched Games: d-prime = 0.20

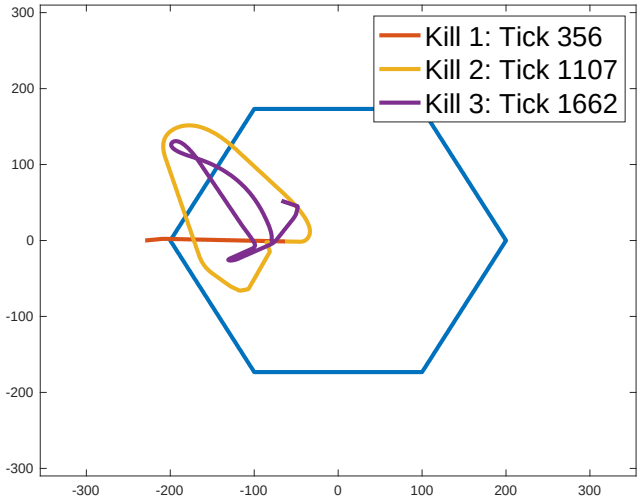
		Marker Present	Maker Absent
Actual Games	Marker Present	25,820	39,932
	Marker Absent	29,450	63,198

- The major reason for the less-than perfect matches is the combinatorics of leg placement which makes 200,000+ model runs not enough.
- However, a lesser reason is the model's failure to produce very short crossing legs and very long turns.

Key Presses?

10^1	10 Seconds	Unit Task	Cognitive
10^0	1 Second	Operations	
10^{-1}	100 Milliseconds	Deliberate act	

- The key presses correspond to some of the Deliberate acts in the model.
- The stitching processes uses the key presses (turn-down, turn-up, thrust-down, thrust-up) associated with a critical segment.
- **Problem:** While ACT-R could, the model does not overlap keypresses. Subjects will sometimes simultaneously press the turn-key and the thrust-key which results in curved paths. Some subject do it hardly ever while others do it a lot.



Conclusion: Limited Results for Model Tracing of Open-Ended Tasks

Scale	Time units	System	World
10^7	Months		Social
10^6	Weeks		
10^5	Days		
10^4	Hours	Task	Rational
10^3	10 Minutes	Task	
10^2	1 Minute	Task	
10^1	10 Seconds		
10^0	1 Second		
10^{-1}	100 Milliseconds		
10^{-2}	10 Milliseconds	Neural circuit	Biological
10^{-3}	1 Millisecond	Neuron	
10^{-4}	100 Microsecond	Organelle	

Unit Task: Stitching works
Operations: Overwhelmed by Combinatorics
Deliberate Act: Model mismatch with Subjects