Strategies, Heuristics and Biases in the Ball Weighing Experiment

Frédéric Dandurand

Seminar in Psychology – McGill University

October 31, 2006

Ph.D. Research

- Problem Solving
- How humans solve problems
- Computational model of problem solving

Problem Solving Task

- Ball weighing experiment
  - (Gizmo problem solving task)

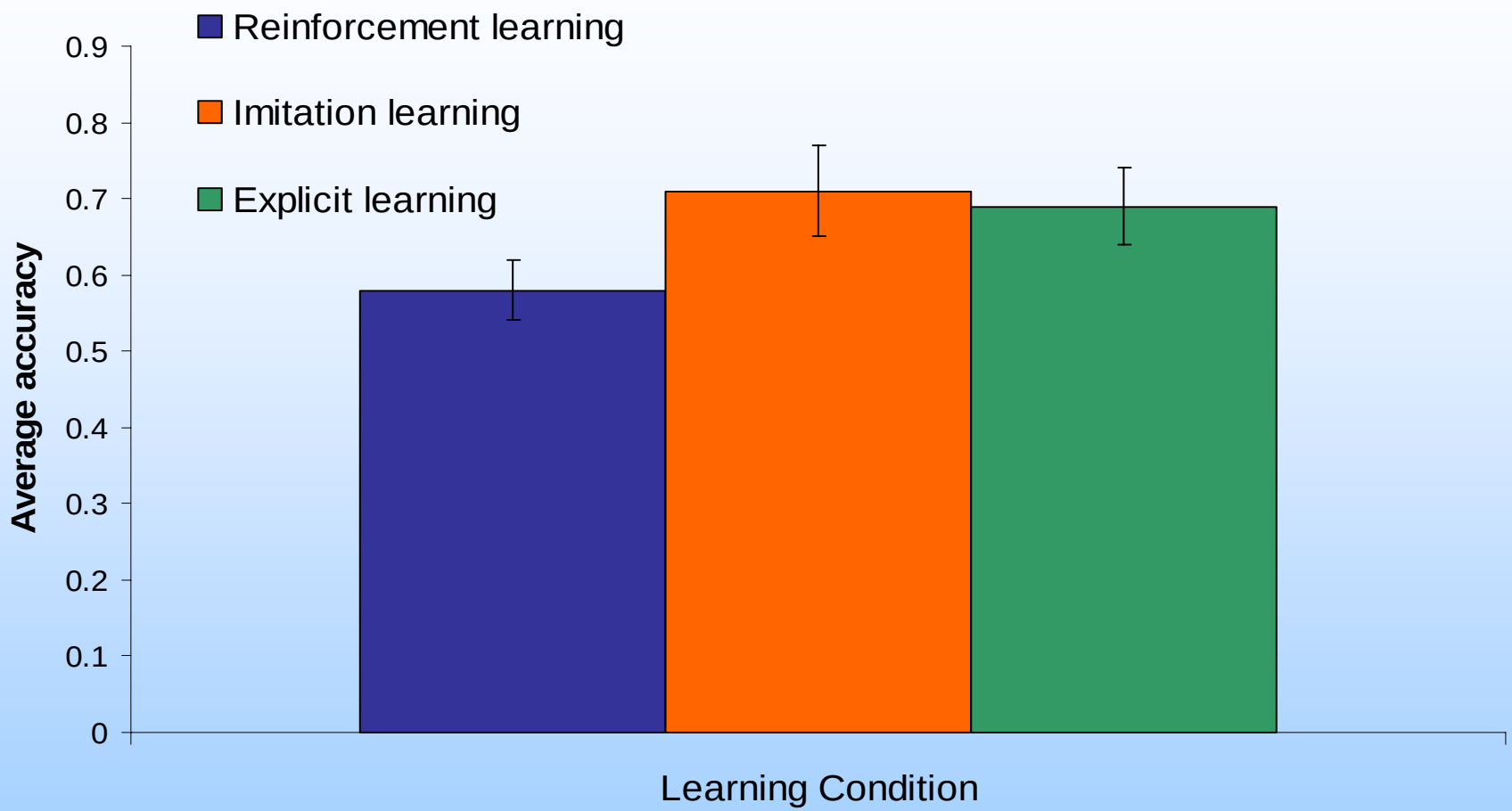
- Find, with **three uses of a scale**, the one ball (gizmo) that is either **heavier** or **lighter** than the rest of a set of 12 balls (gizmo).

- Demos!

Learning Groups

- Study teaching methods (Dandurand, Bowen, Shultz, 2004)
  - Reinforcement learning
    - Got feedback
  - Imitation learning
    - Watched demos
  - Explicit learning (verbal rules)
    - Read instructions

Accuracy on the Ball Weighing Experiment



Imitation and Explicit learning outperform Reinforcement

Today's talk

- Instructions improve performance
- What changed?
  - Strategies, biases, heuristics, etc.
  - Detailed analysis of steps
- Why?
  - Memorizing or understanding
  - Think Aloud Protocols

What changed

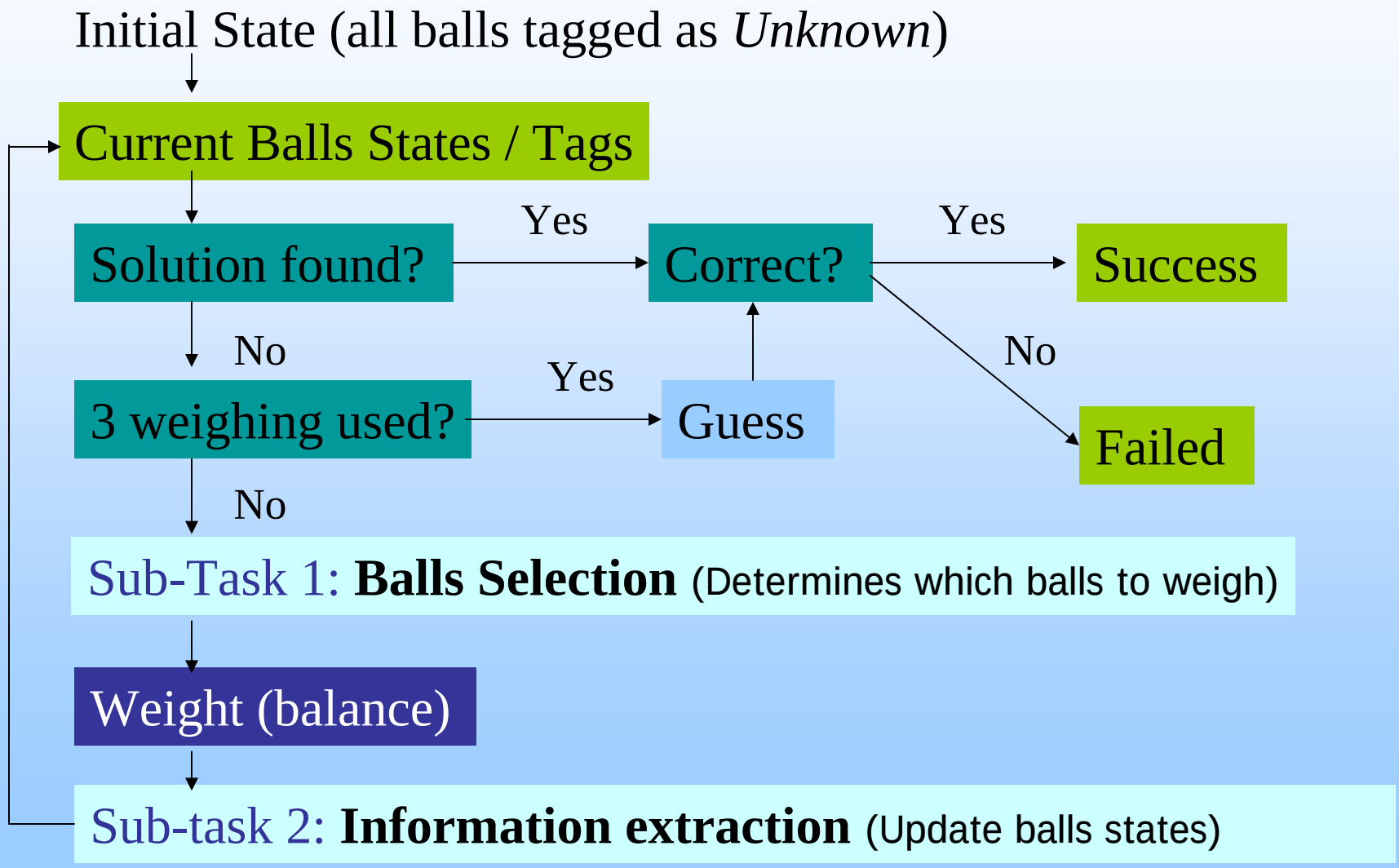
- Task analysis
- Optimal solution
  - Details of steps
  - Characteristics
- Micro-analysis of data
  - Asymmetry and complexity indexes

Task Difficulty

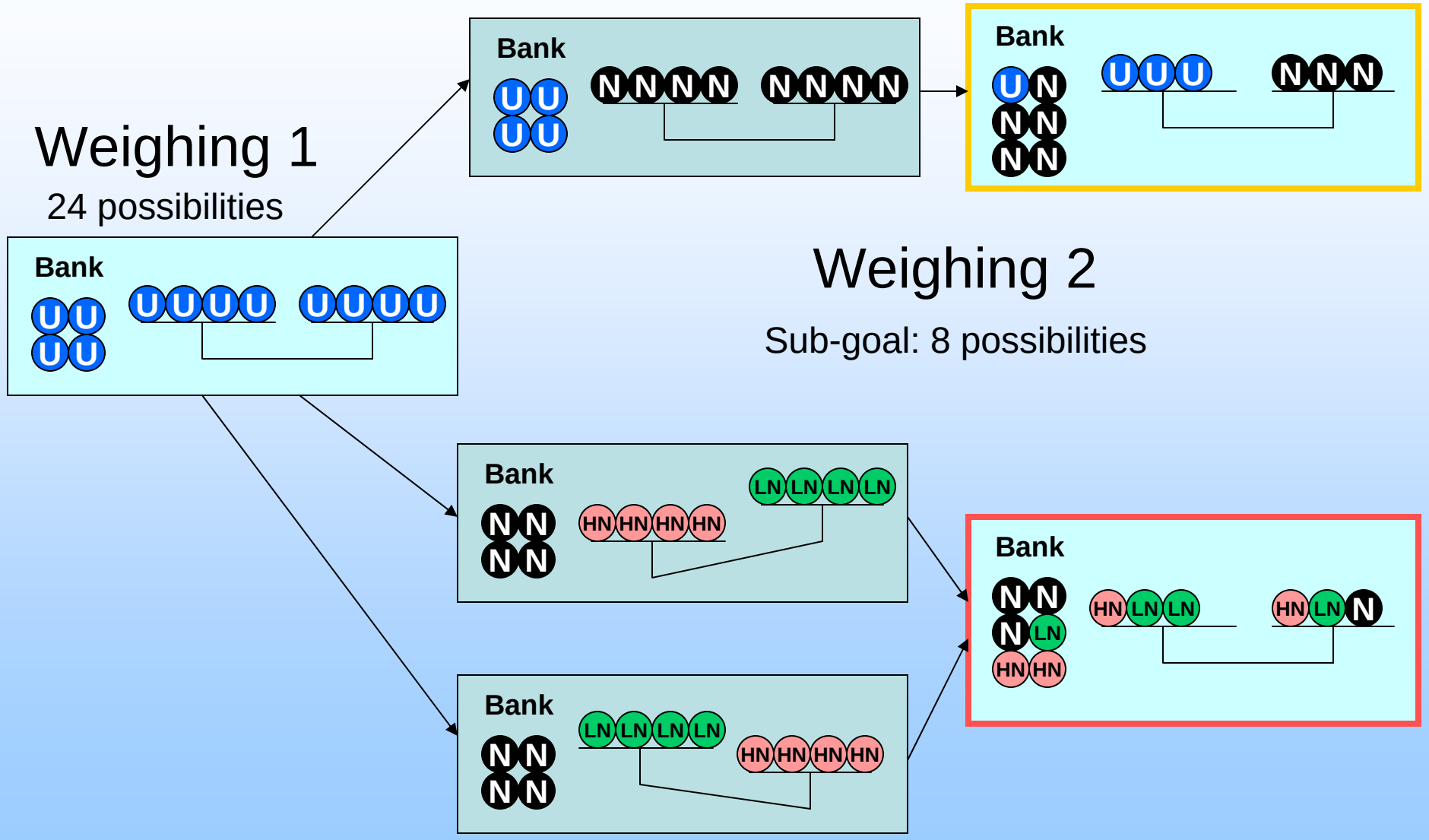
- “It's impossible unless you get lucky.” (Participant 11)

- “I can't even imagine how I’m going to do this.” (Participant 13)

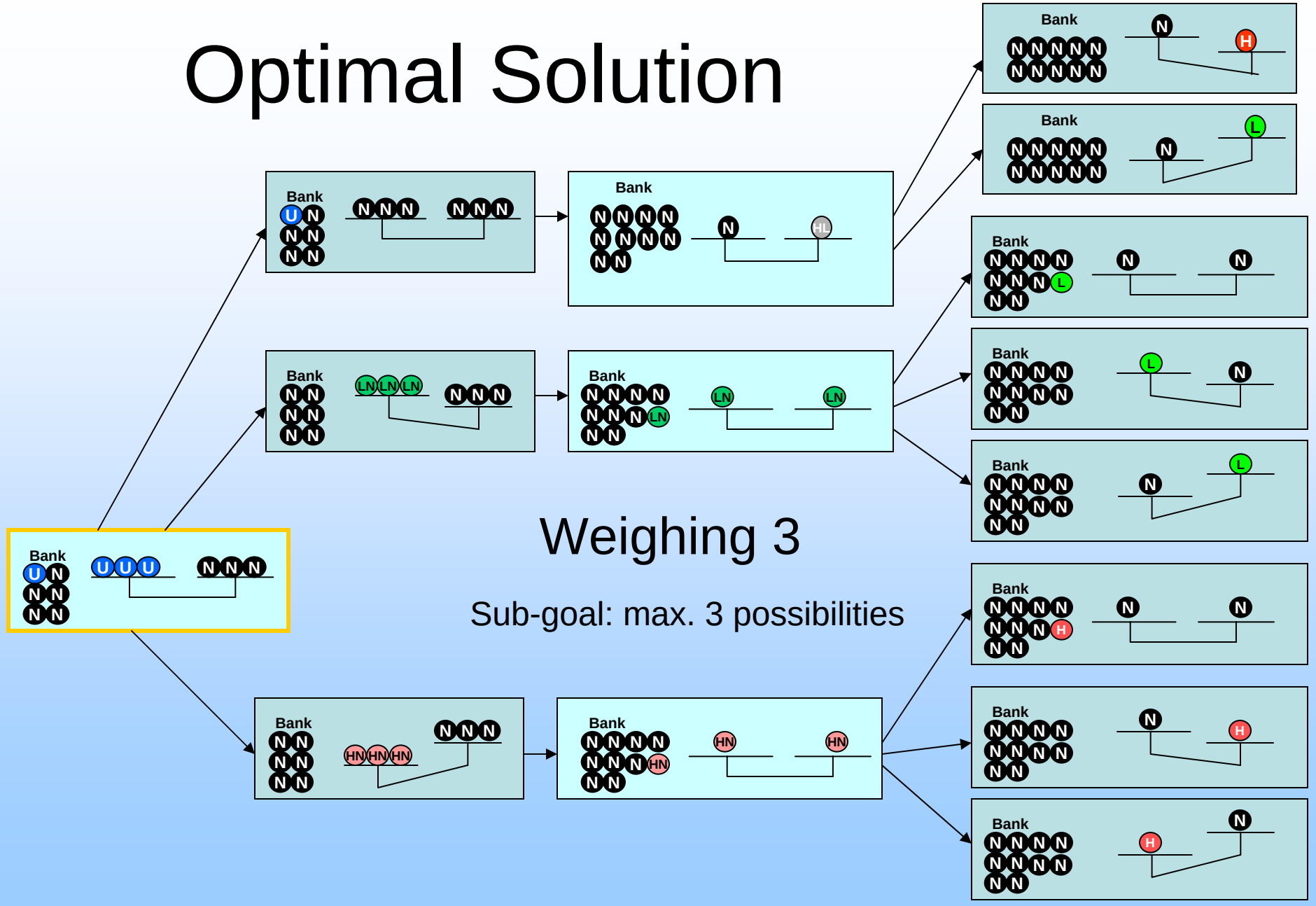
Task Analysis



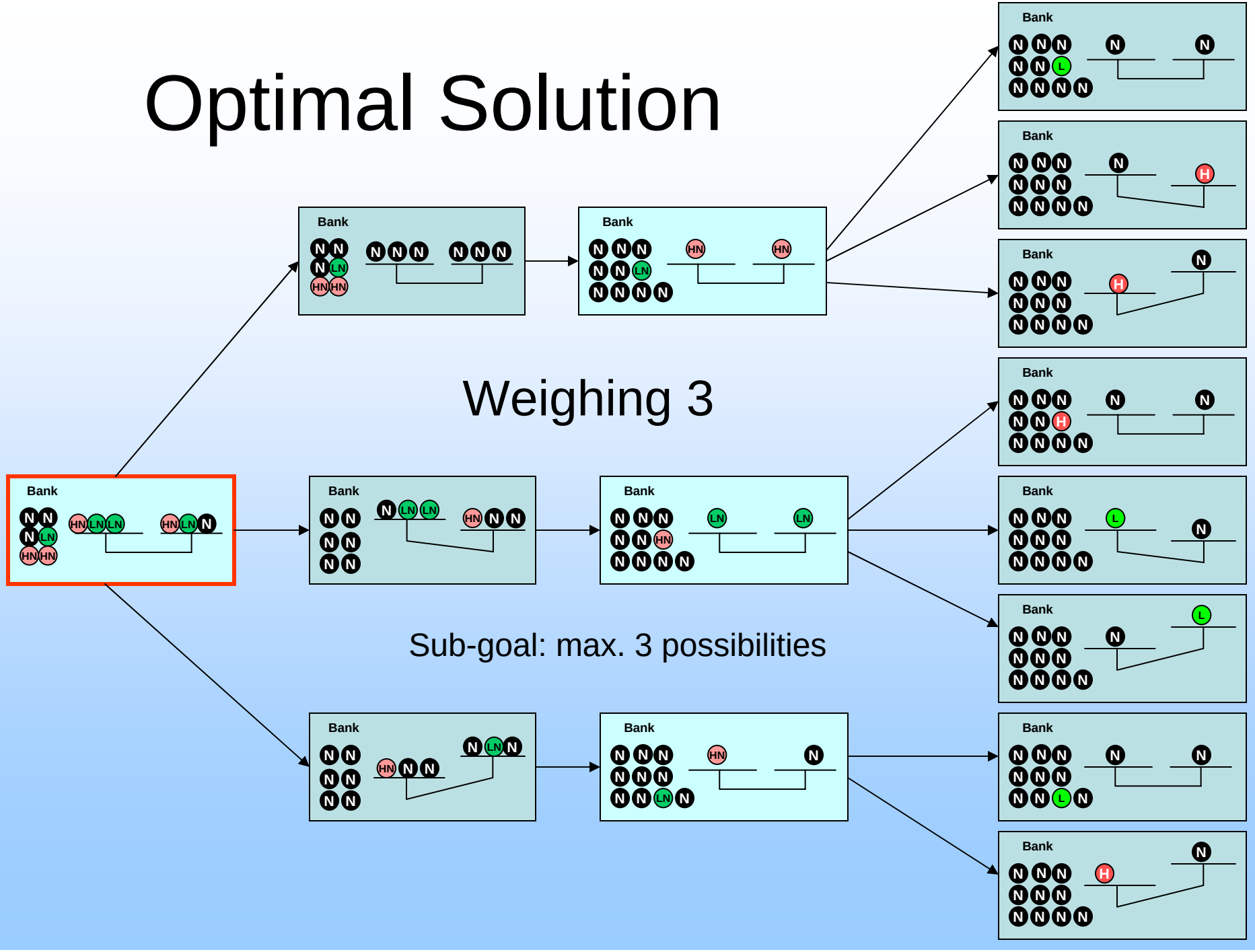
Optimal Solution



Optimal Solution



Optimal Solution



Characteristics of Optimal Solution

- Balls selection
  - Complex – mix balls types
  - Asymmetric – not the same arrangement
- Information extraction (state update)
  - Complex – many updates at each step
  - Attend to the bank
    - Indirect information

Ball Selection Examples

	Complexity	Asymmetry
Items (tokens, balls)	(balls on scale) HN HN vs. N N → 4 LN LN LN vs. LN LN LN → 6 HN LN LN vs. HN LN N → 6	(balls diff. left right) HN HN vs. N N → 4 LN LN LN vs. LN LN LN → 0 HN LN LN vs. HN LN N → 2
Labels (tags) (less redundancy)	(labels on scale) HN HN vs. N N → 2 LN LN LN vs. LN LN LN → 2 HN LN LN vs. HN LN N → 5	(labels diff. left right) HN HN vs. N N → 2 LN LN LN vs. LN LN LN → 0 HN LN LN vs. HN LN N → 1

Note: involves scale only (left and right sides)

Information Extraction Examples

	Before	After	Complexity Index
Items (tokens, balls)	Bank: U U U U Scale: U U U U -U U U U	Bank: U U U U Scale: HN HN HN HN – LN LN LN LN	8
	Bank: U U U U Scale U U U U - U U U U	Bank: N N N N Scale: HN HN HN HN – LN LN LN LN	12
	Bank: HN HN LN N N N Scale: HN LN LN – HN LN N	Bank: N N N N N N Scale: HN N N – N LN N	6
Labels (tags)	Bank: U U U U Scale: U U U U -U U U U	Bank: U U U U Scale: HN HN HN HN – LN LN LN LN	2
	Bank: U U U U Scale U U U U - U U U U	Bank: N N N N Scale: HN HN HN HN – LN LN LN LN	3
	Bank: HN HN LN N N N Scale: HN LN LN – HN LN N	Bank: N N N N N N Scale: HN N N – N LN N	4

Weighing 1 Indexes

	Reinforcement	Imitation	Explicit
Balls selection			
Complexity labels (correct & error)	2.0 & 2.0	2.0 & 1.98	2.0 & 1.95
Asymmetry labels (correct & error)	0.07 & 0.07	0.0 & 0.0	0.01 & 0.0
Information extraction (state update)			
Complexity labels (correct & error)	3.00 & 2.91	3.00 & 2.94	3.01 & 2.98

- No significant difference
  - Balls selection complexity
  - Balls selection asymmetry
  - Information extraction complexity

Weighing 2 Indexes

	Reinforcement	Imitation	Explicit
Ball selection			
Complexity labels	2.25 & 2.3	3.96 & 3.76	3.39 & 3.17
Asymmetry labels	1.18 & 0.87	1.12 & 0.82	1.05 & 0.81
Information extraction (state update)			
Complexity labels	3.75 & 3.74	4.44 & 4.43	4.21 & 4.44

- Balls selection
  - Complexity (Imitation & Explicit) > Reinforcement
  - Asymmetry No main effect, no interaction
- Information extraction complexity

  (Imitation & Explicit) > Reinforcement
- **Conclusion**: Instructions → more complex selections and information extraction (updates)

Weighing 3 Indexes

	Reinforcement	Imitation	Explicit
Ball selection			
Complexity labels	2.31 & 2.43	2.35 & 2.85	2.34 & 2.83
Asymmetry labels	0.78 & 0.92	0.59 & 1.26	0.52 & 0.78
Information extraction (state update)			
Complexity labels	3.52 & 3.88	3.48 & 3.8	3.46 & 3.86

- Balls selection
  - Complexity: Error > Correct
  - Asymmetry: Error > Correct
- Information extraction complexity Error > Correct
- Optimal solution: little variance left in weighing 3
- **Conclusion**: More opportunity left at weighing 3 for complex and asymmetrical steps and updates on erroneous solutions

Attending to the Bank

Bank updates (correct & error)	Reinforcement	Imitation	Explicit	Group	Correct	Inter.
Weighing 1	0.32 & 0.24	0.62 & 0.48	0.51 & 0.34	X	X	X
Weighing 2	0.39 & 0.29	0.52 & 0.34	0.55 & 0.27	X		
Weighing 3	0.74 & 0.8	0.85 & 0.73	0.87 & 0.83		X	X

- Indirect information extracted from the bank
- More attention to the bank
  - Instructions groups
  - Correct trials
- Conclusion: Instructions help participants attend to the bank

Strategies

Strategy	Reinforcement	Imitation	Explicit
3/3	22.9%	0.0%	10.3%
4/4	31.5%	**100%**	**83.3%**
6/6	**34.9%**	0.0%	3.2%
Other	10.7%	0.0%	3.2%

- Reinforcement
  - Prefer 6/6
  - More variance and exploration
- Instructions directly choose 4/4 (correct strategy)
  - Explicit misinterpret 1/3 of 12 as 3/3 (about 10%)

Summary

- Instructions…
  - Reduce simplicity bias
  - Help participants attend to the bank
  - Show correct strategies

Why instructions help?

- Two opposing views
  - Traditional problem solving
    - Learning by demonstration is “rote memorizing” (Katona, 1940)
  - Imitation learning
    - Learning complex hierarchical structures (Byrne & Russon, 1998)
    - Understanding of intentions of the mentor
- So… memorizing or understanding?

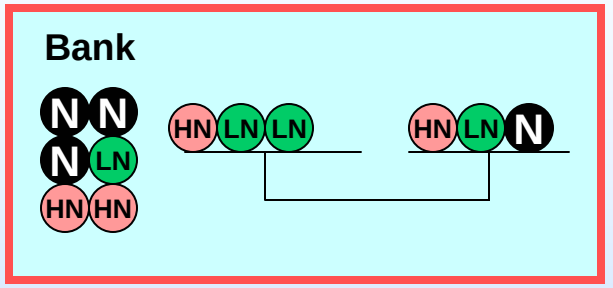
Think Aloud Protocols

- Participants verbalize as they solve problems
- Limitation and warning
  - Access working-memory based processes, but not unconscious ones
  - Reports (especially retrospective) can be inaccurate

Evidence for Memory

Evidence for memory

- More difficult step
  → more memorized



  - "I remember there's a lot of mixing up between known potentially lighter and potentially heavier... I don't remember the exact distribution." (Participant 5)
  - (Participant installs HN LN LN - HN LN N) "is that right? Let's hope so, because if it's heavier... <pause> I don't remember." (Participant 10)

Evidence for memory

- Remembering what was done before
  - “there's only half the chance I will get ... What did I do last time?” (Participant 12)
  - “…same as last time…” (Participant 6)
- Strategies to help remembering
  - "I'm not thinking anything, so I'm going to visually plot some balls and see if that tweaks my memory..." (Participant 5)
  - "Clear the scale first, take a fresh look at this" (Participant 5)

Evidence for Understanding

Advanced Encodings

- Attempt to find an isomorphic representation (mathematical)
  - “if this was numbers, it would be easier. Unidentified could be ones, lighters could be negative one, and the heavier one could be two. Ok, mathematical strategy.“

    (Participant 4)
  - “If I had some paper, I could write some equations, that might help!” (Participant 5)

Explicitly Identifying Sub-goals

- Have 3 possibilities at most at 3rd weighing
  - "You have to make sure that after 2 steps, you are left with only 3 balls that might not be of normal weight" (Participant 2)
  - "Somehow in this test, I have to like reduce my unknown balls to 3, and then I can solve this." (Participant 5)

Providing Explanations

- Explicit (verbal rules) group
  - "If they're light or normal ones that move down, then they are normal. Yeah, because they're obviously not light, they went down" (Participant 10)
- Imitation Group
  - Reason and explain state update in the bank

Generalization

- Identify general and regular patterns
  - Ex: Weighing 1: 4/4 W2: 3/3 W3: 1/1
  - Difficult because few regular patterns
- Recognize certain configurations as similar to others
  - Weighing 3: HN HN LN similar to LN LN HN

Example

- Participant 2 11:00 – 12:00
  - Explain state update in the bank
  - Recognize weighing 3 as similar to previous one

Example

- Participant 5 17:43-18:54
  - General patterns (4-3-1)
  - Understanding ball markings

Evidence for Reasoning

Individual Differences

- Low
  - “You're not really sure why you're doing it a certain way… I just click, it's automatic.” (Participant 0, retrospective report)
- High
  - “So I weigh the two heavy ones, and they're the same, then I know it's the light one different, and if one of them is heavier than the other, then I know it's heavy. Aahh! I figured it out.” (Participant 10)

Example

- Participant 5 37:18 – 38:40
  - Use of reasoning to reconstruct the correct step
  - Not everything is explicit

Integration of Cognitive Mechanisms

Follow Instructions or Explore?

- Participant 10 asked the experimenter if she can do it any other way than what was shown in the instructions. She said she was getting bored.
- "Right now I'm thinking if there is any different way to solve this problem, besides the instructions I was given.“ (Participant 5)
- "I wonder... what if I put two on one <pause> Mmm. That's for later. Stick to the strategy I have been using“ (Participant 6)

Integration of Memory and Reasoning

- Considering cases not covered in instructions
  - “…so if it moves after that, one LN from the scale vs. one LN from the scale. But what if it was a heavy one that was different. Would that not work?"

    (Participant 10)

- Fill-in missing memories using reasoning
  - "I don't quite remember those instructions, so I'm going to have to think a bit on what to do next.” (Participant 5)

Choosing between Memory and Reasoning

- Conflict between incorrect memories and reasoning (Participant 9)
  - Trial 7: Reasoning 52:38 – 53:30
  - Trial 8: Memory 56:15 – 57:00

Example

- Participant 2: 15:38 – 19:02
  - Choosing between demo and exploration
  - Experimenting with constraints: checking if program allows more than 4 balls per side
  - Memory fault: thinks he has not done as the demo, but he actually did
  - Correct reasoning
  - Self-generate feedback: no explicit feedback but he knows he's correct

Memorizing or Understanding?

- Evidence for both
- Understanding
  - Advanced encodings
  - Identifying sub-goals
  - Providing explanations
  - Generalization (regularity, similarity)
  - Reasoning
- Memorizing
  - Difficult steps
- Individual differences

Take Home Messages

- Watching demonstrations…
  - Complex processes involved
  - Clearly not rote memorizing
- Individual differences

Future Work

- Role of motivation
- Computational model

The End

- Acknowledgements
  - Bernardo A. M. de Alvarenga
  - Thomas R. Shultz
  - Kristine H. Onishi
- Thank you for your attention!
- Questions or comments?