UF Herbert Wertheim College of Engineering
UNIVERSITY of FLORIDA

Designing for Children to Build Conversational Agents and Learn about Artificial Intelligence

Xiaoyi Tian

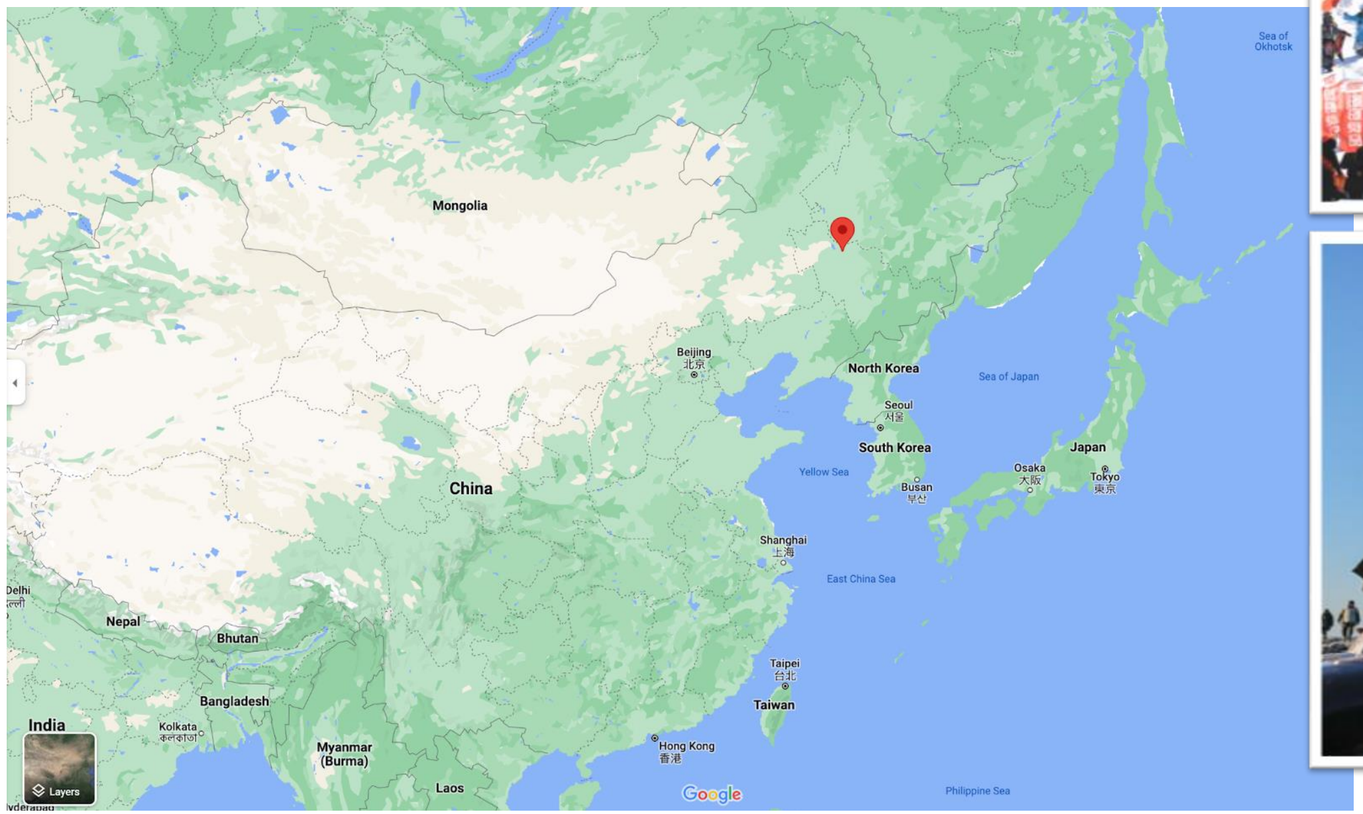
Ph.D. Defense

Committee Members: Kristy Boyer (Chair), Eric Ragan, Jaime Ruiz, Maya Israel

5/28/2024

About me

From Songyuan, northeast China



About me

B.S. in Management Science, 2018

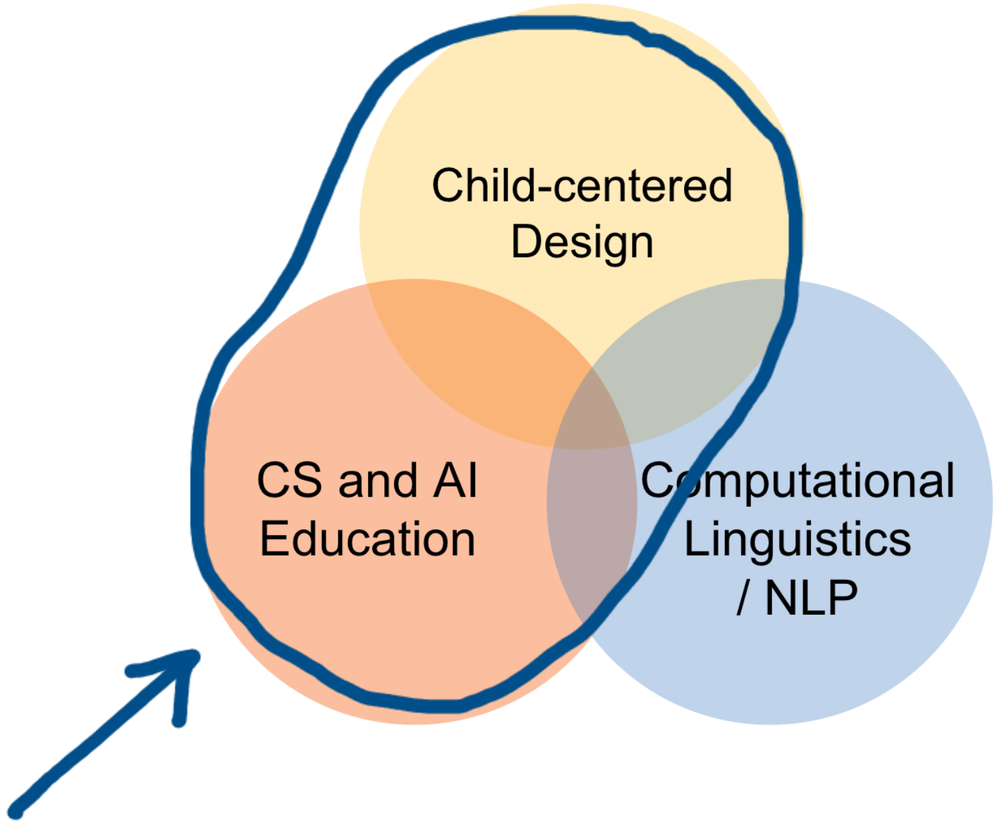
Anhui University, China

M.S. in Information Science, 2020

University of Pittsburgh

Ph.D. in Human-Centered Computing, 2024

University of Florida



My dissertation

Introduction

- Conversational AI is ubiquitous in everyday life
  - Siri, Alexa, and ChatGPT
- However, children are not afforded to ***understand*** the inner workings of AI
- Teaching young students about AI

How can we provide engaging and authentic AI learning experiences for children?

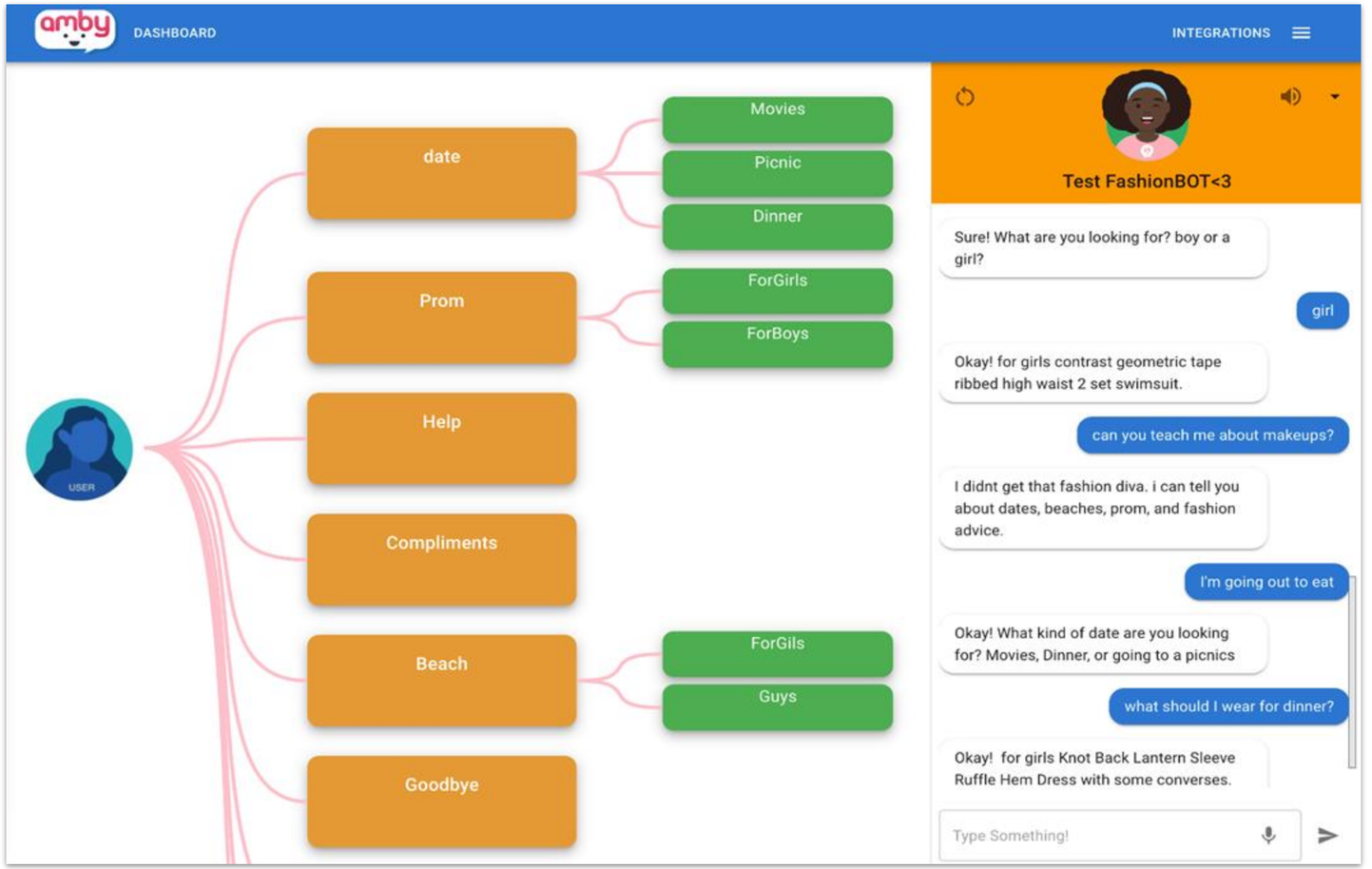
Theories of Constructionism and Authenticity

"*The best way to do that is to build something tangible – something outside your head – that is also personally meaningful.*" – Seymour Papert (1987)

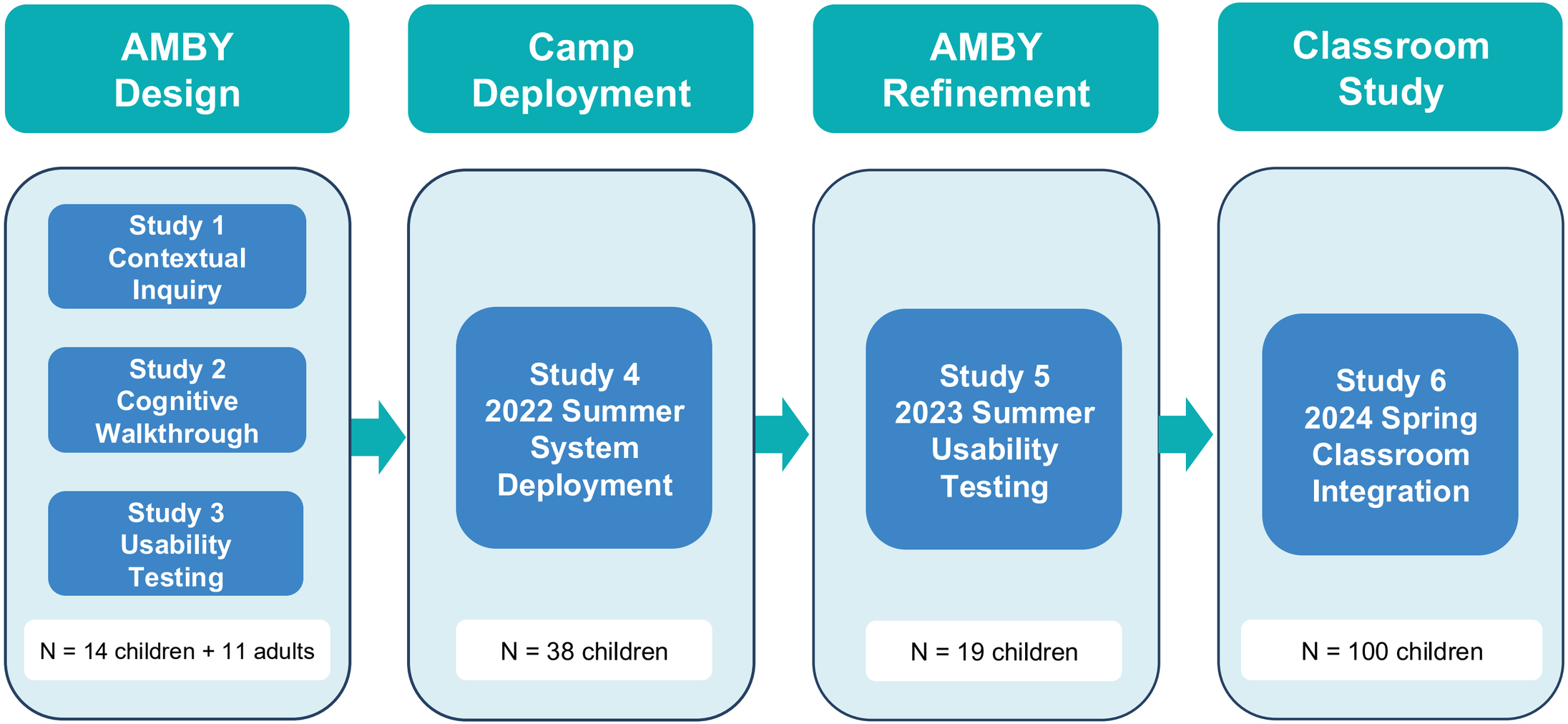
- Constructionism-based applications are prevalent in computing education
  - E.g., Scratch, App inventor
- Our approach: introducing AI through building conversational agents
- Learning environment with "Thick Authenticity" (Shaffer and Resnick, 1999)
  - Personal
  - Real-world
  - Discipline-specific
  - Assessment

AMBY - Conversational AI development platform

"AI Made By You"



Dissertation Study Overview

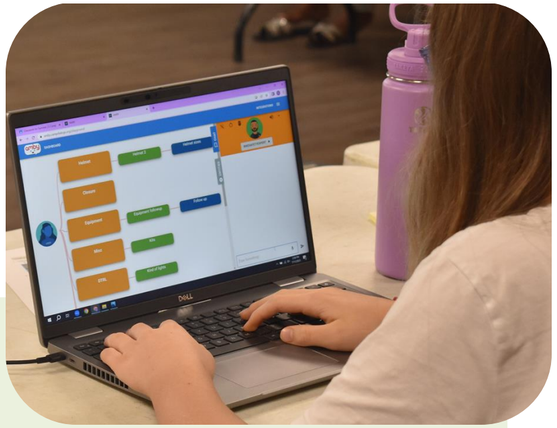


Phase 1: AMBY 1.0 Design

AMBY Design

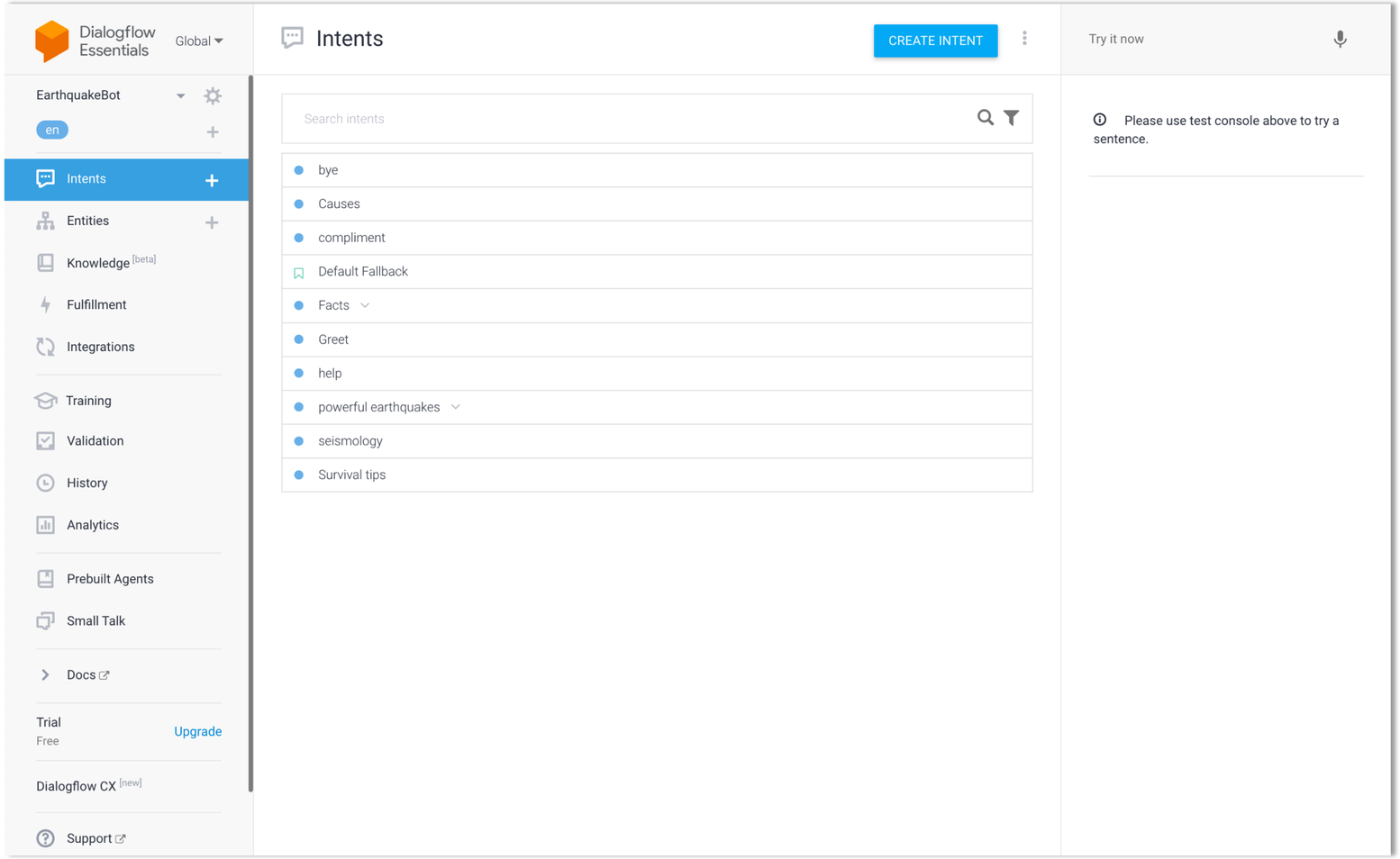
- **Study 1** Contextual Inquiry
- **Study 2** Cognitive Walkthrough
- **Study 3** Usability Testing

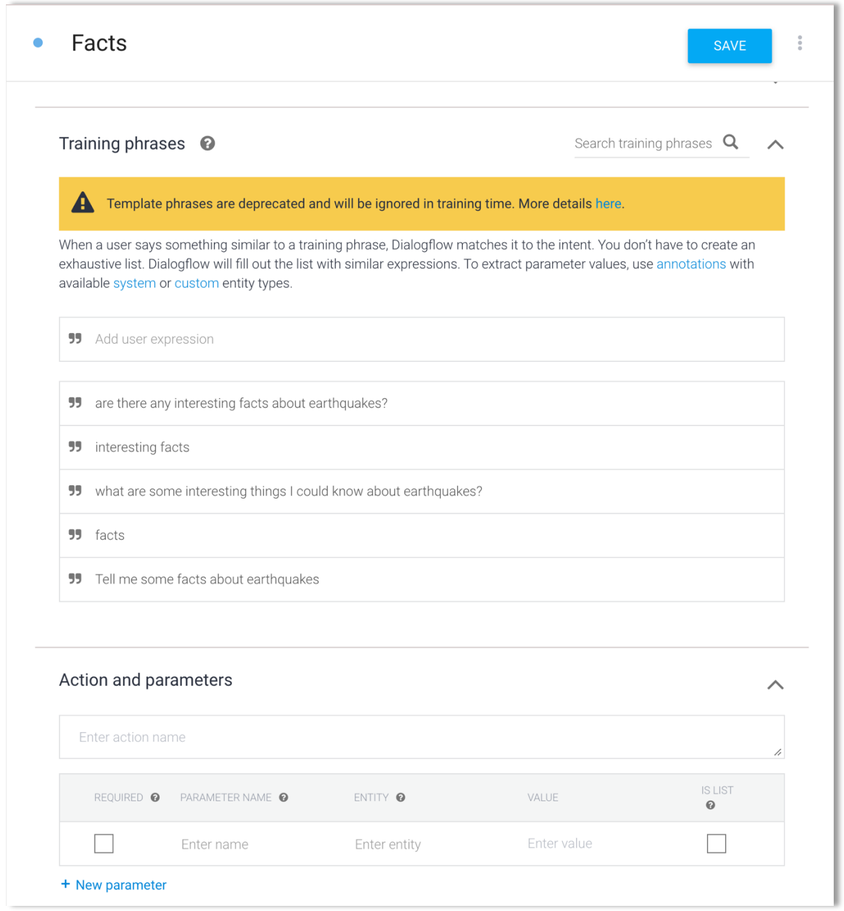
N = 14 children + 11 adults



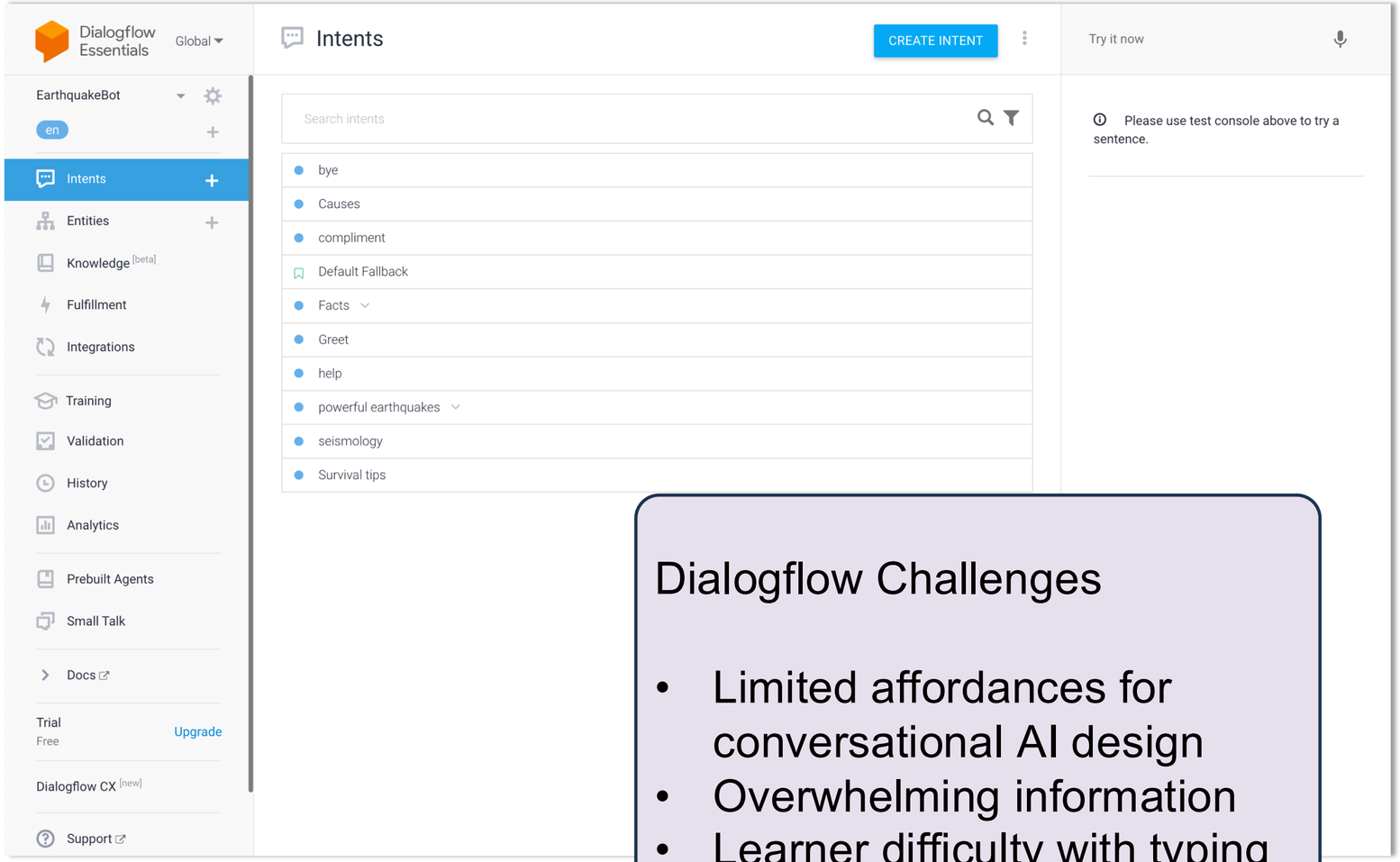
Design goal:
A learning environment that empowers children with conversational AI development

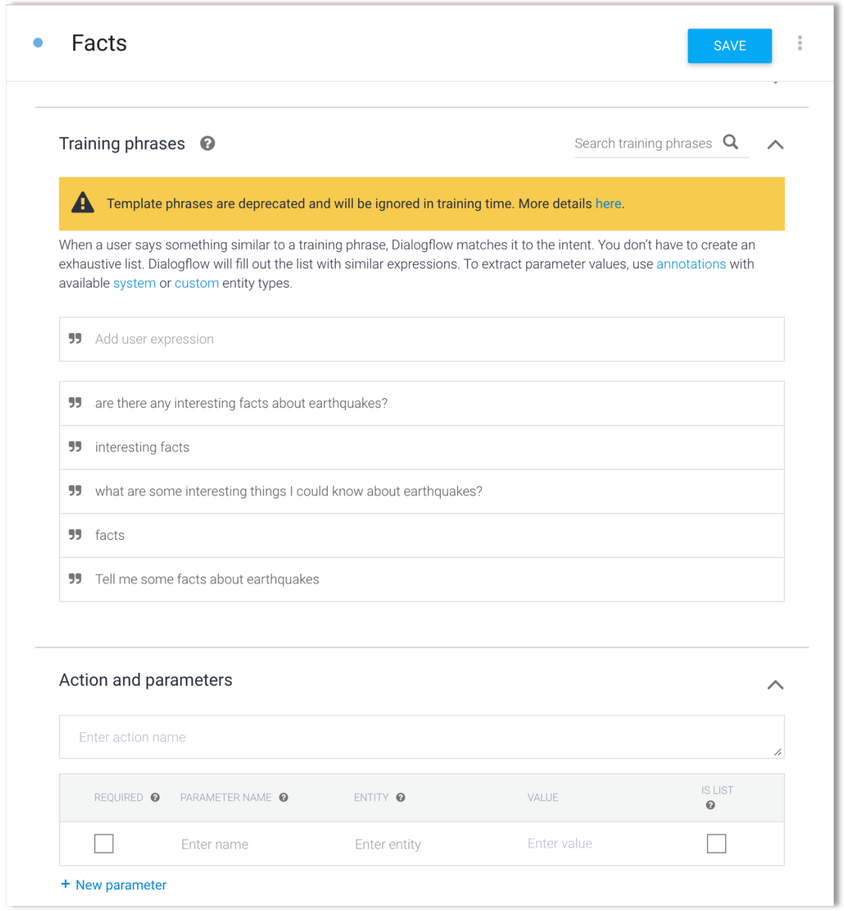
Where We Started: Google Dialogflow Interface





Where We Started: Google Dialogflow Interface

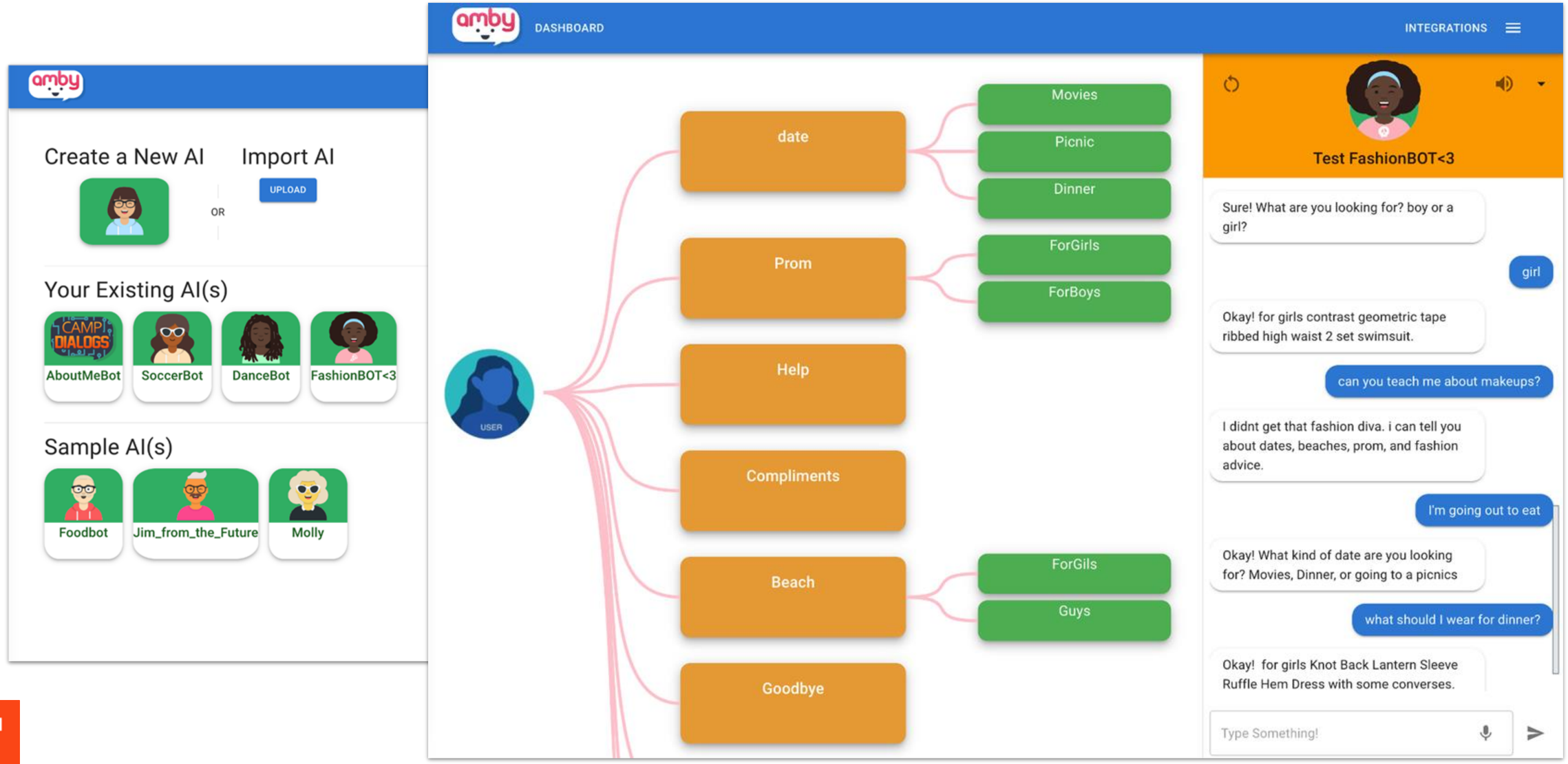




Dialogflow Challenges

- Limited affordances for conversational AI design
- Overwhelming information
- Learner difficulty with typing

Our New Interface: AMBY

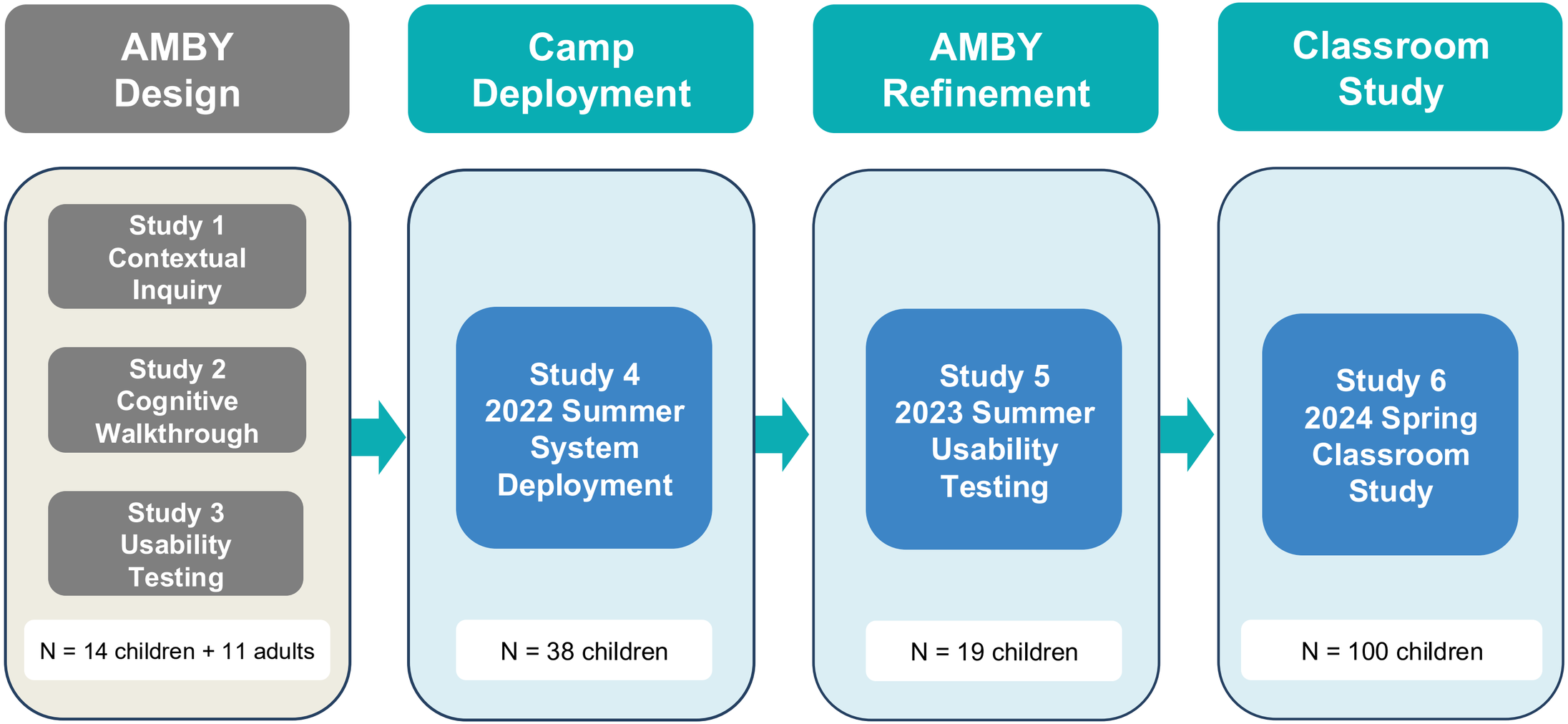


AMBY Demo

YouTube Link

I would like to acknowledge the software developers who contributed to the AMBY codebase, including Amit Kumar, Sunny Dhama, John Tran Hoang, as well as the scholars who shaped its design and ideation, including Gloria Katuka, Yukyeong Song, Mehmet Celepkolu, Lydia Pezzullo, Joanne Barrett, Tom McKlin, Kristy Boyer, and Maya Israel. Special thanks to Carly Solomon for recording the AMBY demo video.

Dissertation Study Overview



Phase 2: AMBY 1.0 Summer 2022 Deployment

Camp Deployment

Study 4
2022 Summer System Deployment

N = 38 children

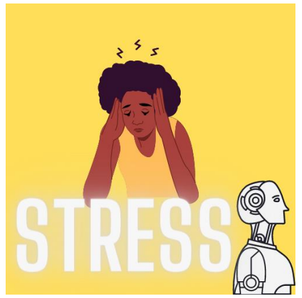
RQ1: How do children engage with a development environment designed to support them in making conversational agents?

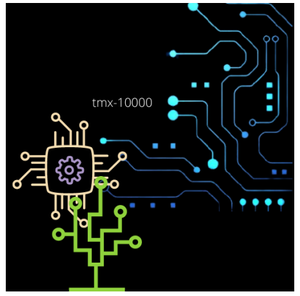
Summer Camp Deployment

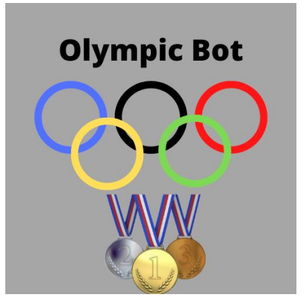
- 38 children* attended the two-week summer camp
- Rising 7th or 8th graders, average age was 12.7 (SD = 0.7)
- 19 girls and 19 boys
- 31 were Black/African American (86%), 5 Hispanic/Latinx, 4 White/Caucasian, 1 Asian and one prefer not to say
- 14 participants (37%) reported having no prior coding experience

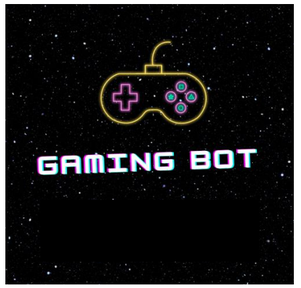
*5 learners were returning participants of 2021 summer camp (study 1)

Chatbots Created by Learners









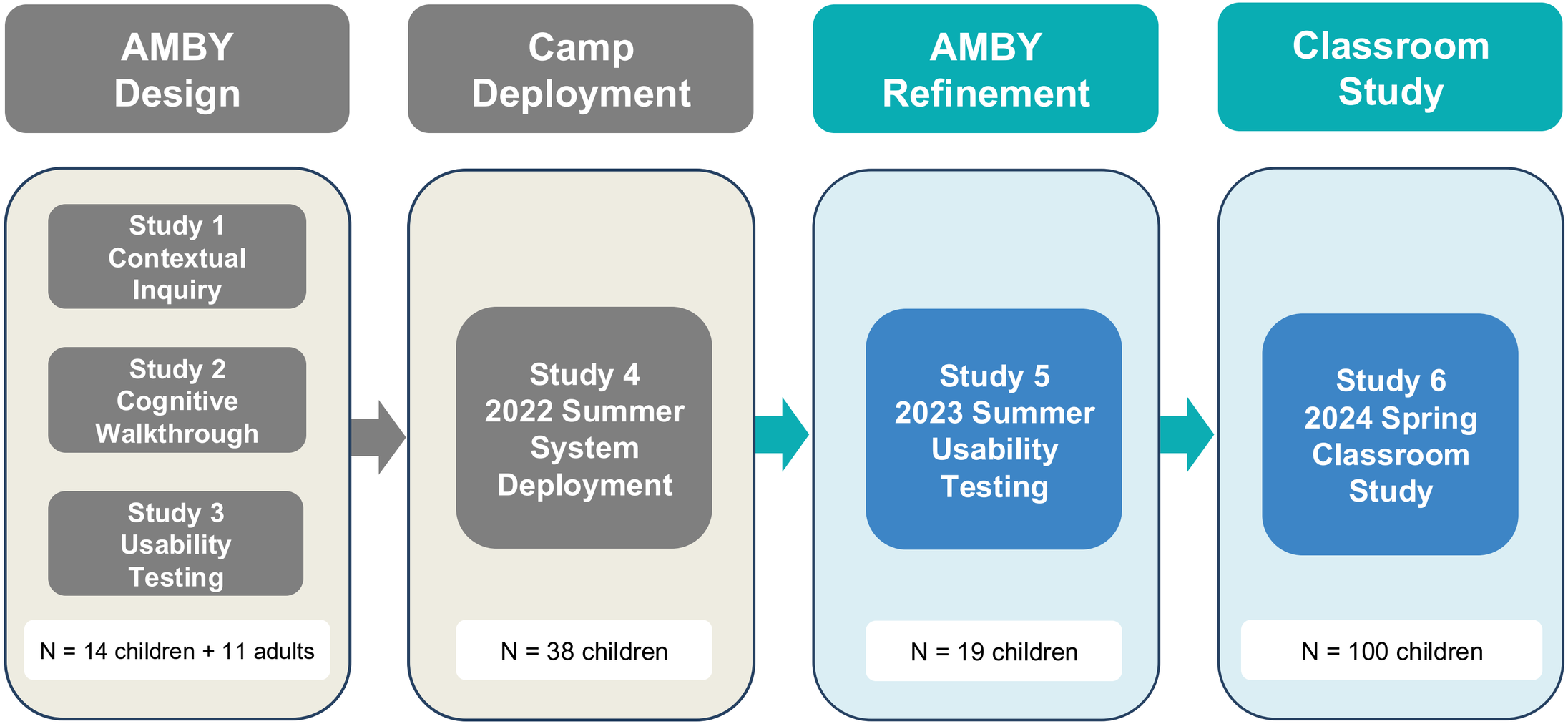
UF UNIVERSITY of FLORIDA

LearnDialogue

Takeaways from Phase 1 & 2 Studies

- Our team created AMBY, a conversational AI development environment for children
- AMBY engaged learners with no prior computing experience, allowing them to create personally meaningful conversational agents with a positive overall experience
- Design lessons learned
- Additional development is needed for the interface to provide “high-ceiling”

Dissertation Study Overview



Phase 3: AMBY Refinement and Summer 2023 Study

AMBY Refinement

Study 5
2023 Summer Usability Testing

N = 19 Children

RQ2: What features do children desire in a learning environment to support their educational needs?

AMBY 2.0: Entity Feature

What are Entities?

Entities **are specific terms that can be extracted from user input and classified into predefined categories.**

What are Entities?

Entities **are specific terms that can be extracted from user input and classified into predefined categories.**

For examples:

"*Tell me the flight information from Orlando to Atlanta.*"

"*Is there a flight from Seattle to Pittsburgh?*"

"*My favorite fruit is apple.*"

"*My favorite fruit is strawberry.*"

Location

Fruits

Entity & Abstraction

- Abstraction is a fundamental concept in computing (Armoni, 2013)
  - Simplify complex systems
  - Hide irrelevant details
  - Focus on the main ideas (Wing, 2008)
- In AMBY, entities help
  - Simplify repetitive training phrases
  - Focus on important tasks
    - E.g., conversation design

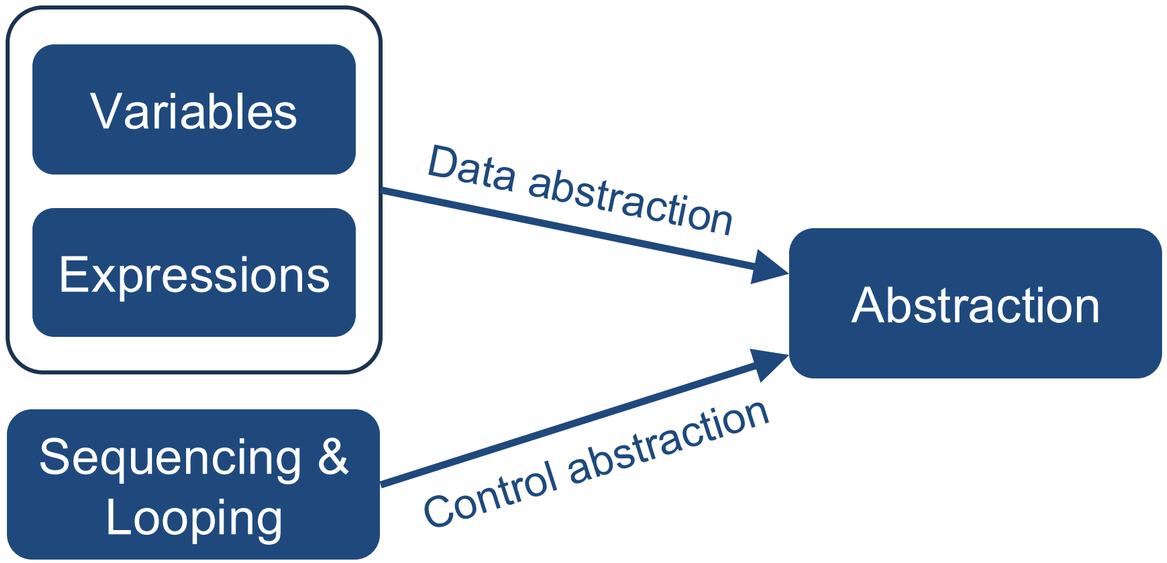


Figure: Abstraction concepts in CS Programming (Grover et al., 2019)

Entity & Abstraction

- Abstraction is a fundamental concept in computing (Armoni, 2013)
  - Simplify complex systems
  - Hide irrelevant details
  - Focus on the main ideas (Wing, 2008)
- In AMBY, entities help
  - Simplify repetitive training phrases
  - Focus on important tasks
    - E.g., conversation design

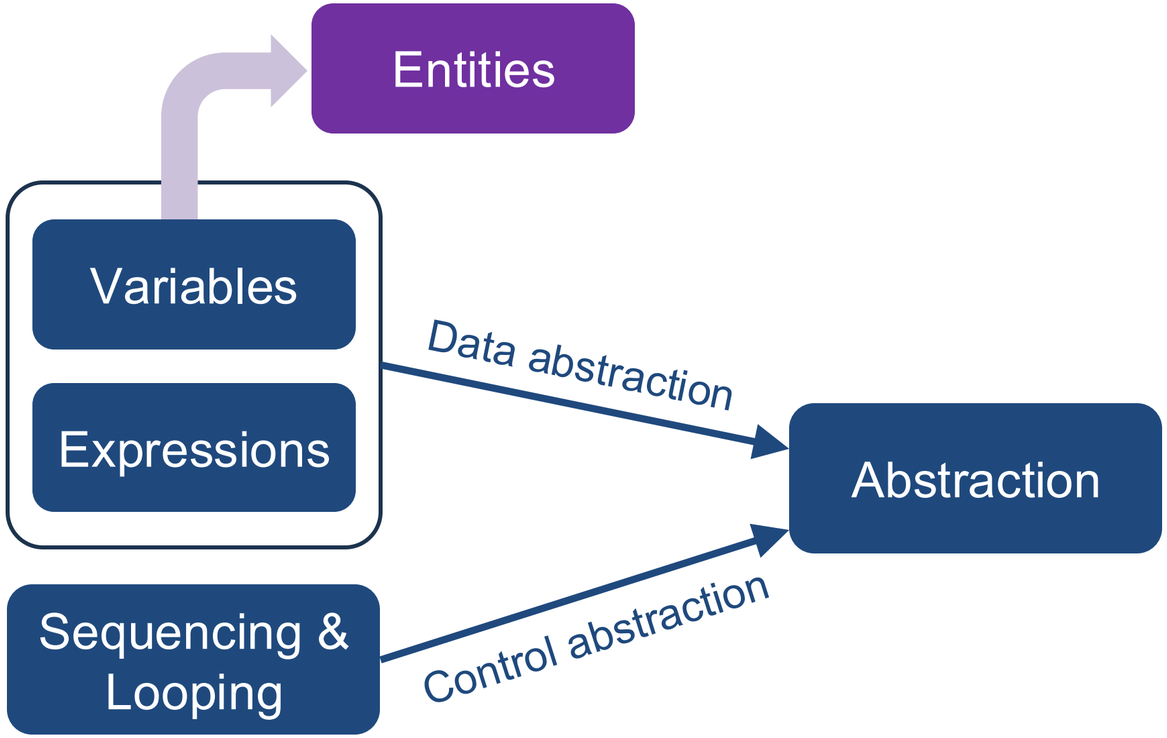


Figure: Abstraction concepts in CS Programming (Grover et al., 2019)

Teaching Abstraction to Children

- Challenges
  - Not sufficiently emphasized by the teachers (Nakar and Armoni, 2023)
  - CS undergraduates' struggles with algorithms (Armoni, 2013)
  - Reluctance to utilize abstract classes in object-oriented programming (Oh-Bach and Lavy, 2004)
- Cognitive development considerations for children (Piaget et al., 1952)
- Start teaching abstraction skills early
- Continually revisit throughout all educational levels (Mirolo et al., 2022)
- *Entities* in AMBY aims to enhance abstraction skills and NLP understanding

Entity Feature Use Case

- **Reduced Training Phrases**

Without Entity, training phrases for “favorite fruit” intent:

I like Apple	Pear is the best	My favorite fruit is apple
I like Kiwi	Mango is the best	My favorite fruit is pear
I like Mango	Kiwi is the best	My favorite fruit is mango
I like Pear	Apple is the best	

- More Personalized Responses

Entity Feature Use Case

Entity: **Fruits**
Juicy -> Mango, Kiwi
Crunchy -> Apple, Pear

- **Reduced Training Phrases**

With Entity, only these training phrases are needed:

I like Apple	Pear is the best	My favorite fruit is apple
~~I like Kiwi~~	~~Mango is the best~~	~~My favorite fruit is pear~~
~~I like Mango~~	~~Kiwi is the best~~	~~My favorite fruit is mango~~
~~I like Pear~~	~~Apple is the best~~	

- More Personalized Responses

Entity Feature Use Case

Entity: **Fruits**
Juicy -> Mango, Kiwi
Crunchy -> Apple, Pear

- Reduced Training Phrases
- **More Personalized Responses**

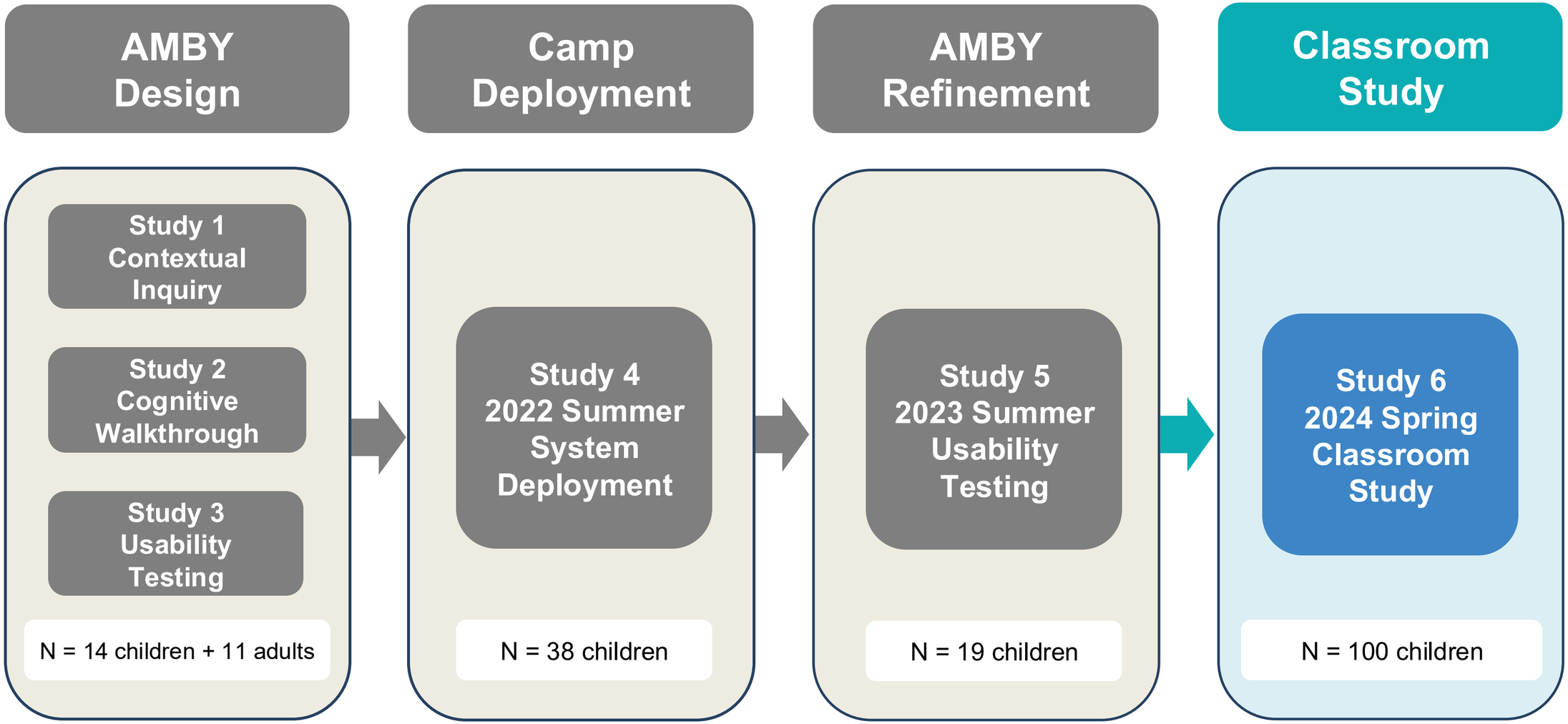
Conversation without entity

Conversation with entity

Takeaways from Phase 3 Usability Testing

- Entity feature shows promise in enhancing learning experience
  - Most participants familiar with entity concept, recognized the utilities of entity
  - Some participants adopted heavily in projects
- Introduce entity concept earlier in curriculum
- Next steps:
  - Assess the entity feature’s effectiveness on a larger scale
  - Deploy AMBY 2.0 in middle school classroom context

Dissertation Study Overview



Phase 4: AMBY 2.0 Classroom Study

Classroom Study

Study 6
2024 Spring Classroom Study

N = 100 children

RQ3: Does the *entity* feature impact students' enjoyment and artifact quality?

Study Design

RQ3: Does the *entity* feature impact students' enjoyment and artifact quality?

- **Experimental Conditions:**
  - Between-subject experiment
  - Two versions of AMBY: AMBY with entity and AMBY without entity
- **Hypothesis:**
  - Hypothesis 1: The entity feature will **enhance students' enjoyment** in creating chatbots, as indicated by the post-questionnaire.
  - Hypothesis 2: The chatbot artifacts produced by students in the entity condition will exhibit **higher project quality**.

Science-Based Conversational AI Curriculum

- Adjust from informal summer camp to classroom settings
  - Condensed the curriculum to a 10-hour learning module
  - Six learning units
- Integrate science content into the curriculum
  - Exemplar Science Chatbots

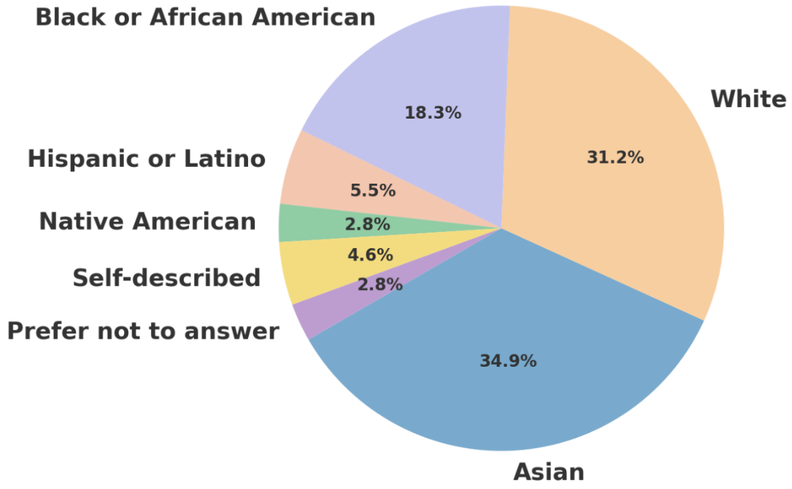
Ralph

ScienceGenius

EarthQuakeBot

Study Context and Participants

- 6th grade Science class
- 128 students in total, 100 consented to participate in research
- Average age: 11.7 (sd = 0.48)
- Gender: 49 girls, 46 boys, 1 non-binary, 1 prefer not to answer
- Race/Ethnicity Distribution



Conditions

Class sections	Participants #
Period 1	15
Period 2	16
Period 3	16
Period 4	16
Period 5	13
Period 6	24

AMBY without entity: N = 47

AMBY with entity: N = 53

Class average grades for all periods were similar prior to the study

Study Procedure

Conditions
- AMBY w Entity
- AMBY w/o Entity

Entity Lesson & Hands-on practice

Pre-survey

Conversational AI Lessons

AMBY Hands-on practice

AI Project Development

Post-survey Post-assessment

Interview

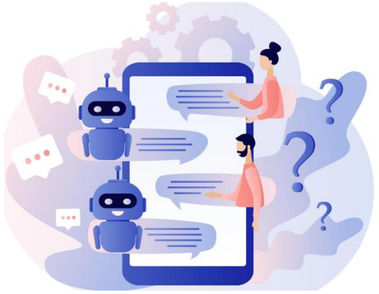
Day 1 | Day 1-3 | Day 4 | Day 5-8 | Day 9 | Day 10

Exception: one class section spent one extra day (50 mins) on the lessons due to the size of the class

Data Collection

Pre-/post-questionnaires

Focus group interview



Student-created chatbot artifacts

Post-assessment

Results: Impact on Enjoyment

- *Enjoyment* statements in post-questionnaire, rated on a 4-point Likert Scale
  1. Creating a chatbot is exciting.
  2. Creating a chatbot is enjoyable.
- No statistically significant difference

Condition	N	Mean	SD	Indep t-test
Non-Entity	46	3.41	0.76	p = 0.837, t = -0.2
Entity	51	3.44	0.57	

Results: Impact on Project Quality

Chatbot Artifact Evaluation Dimensions

1. Project ideation
2. Conversational design
3. AI development

Validated rubric (Cohen's Kappa = 0.751)

4. End-user satisfaction (EUS)
   - Six items adapted from Walker et al. (2000)
   - Averaged from three external annotators

Special thanks to Christine Wise, Joanne Barrett, Yukyeong Song, Amit Kumar, John Hoang for helping with rubric development, and thanks to Carly Solomon, Wesly Menard, David Vallejo-Lozano and Madison Edwards for artifact data annotation.

Results: Impact on Project Quality

	All projects (n=51)		Non-entity (n=24)		Entity (n=27)		P-value
	Mean	SD	Mean	SD	Mean	SD	
Project Ideation	3.04	0.4	3.04	0.46	3.04	0.34	0.97
Conv Design	3.24	0.35	3.26	0.34	3.22	0.35	0.68
AI Development	3.16	0.25	3.19	0.29	3.13	0.22	0.42
EUS	3.45	0.81	3.56	0.72	3.36	0.88	0.38

No significant difference

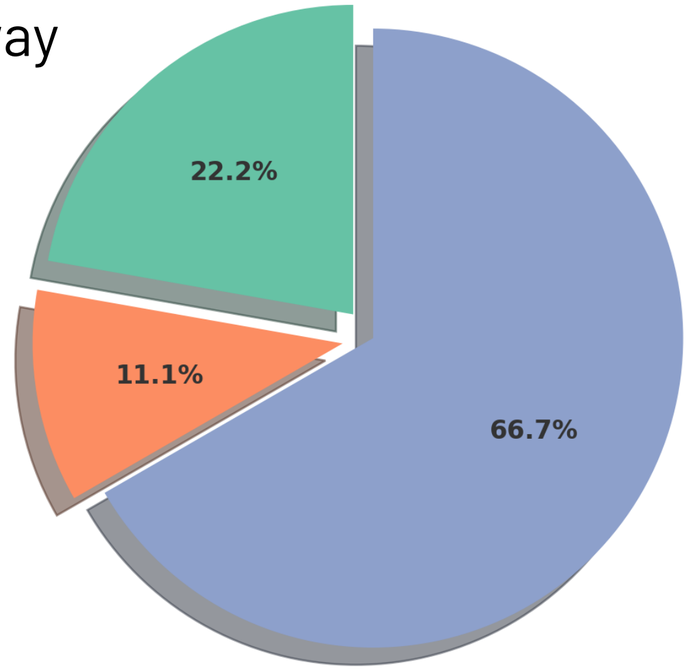
Results: Impact on Project Quality

	All projects (n=51)		Non-entity (n=24)		Entity (n=27)		P-value
	Mean	SD	Mean	SD	Mean	SD	
Project Ideation	3.04	0.4	3.04	0.46	3.04	0.34	0.97
Conv Design	3.24	0.35	3.26	0.34	3.22	0.35	0.68
AI Development	3.16	0.25	3.19	0.29	3.13	0.22	0.42
EUS	3.45	0.81	3.56	0.72	3.36	0.88	0.38

Results: Entity feature usage

Used entity in an insignificant way (6 projects)

Significantly integrated entity (3 projects)



Did not use entity (18 projects)

Entity Feature Usage in 27 Student Projects

Results: 27 Projects in Entity condition

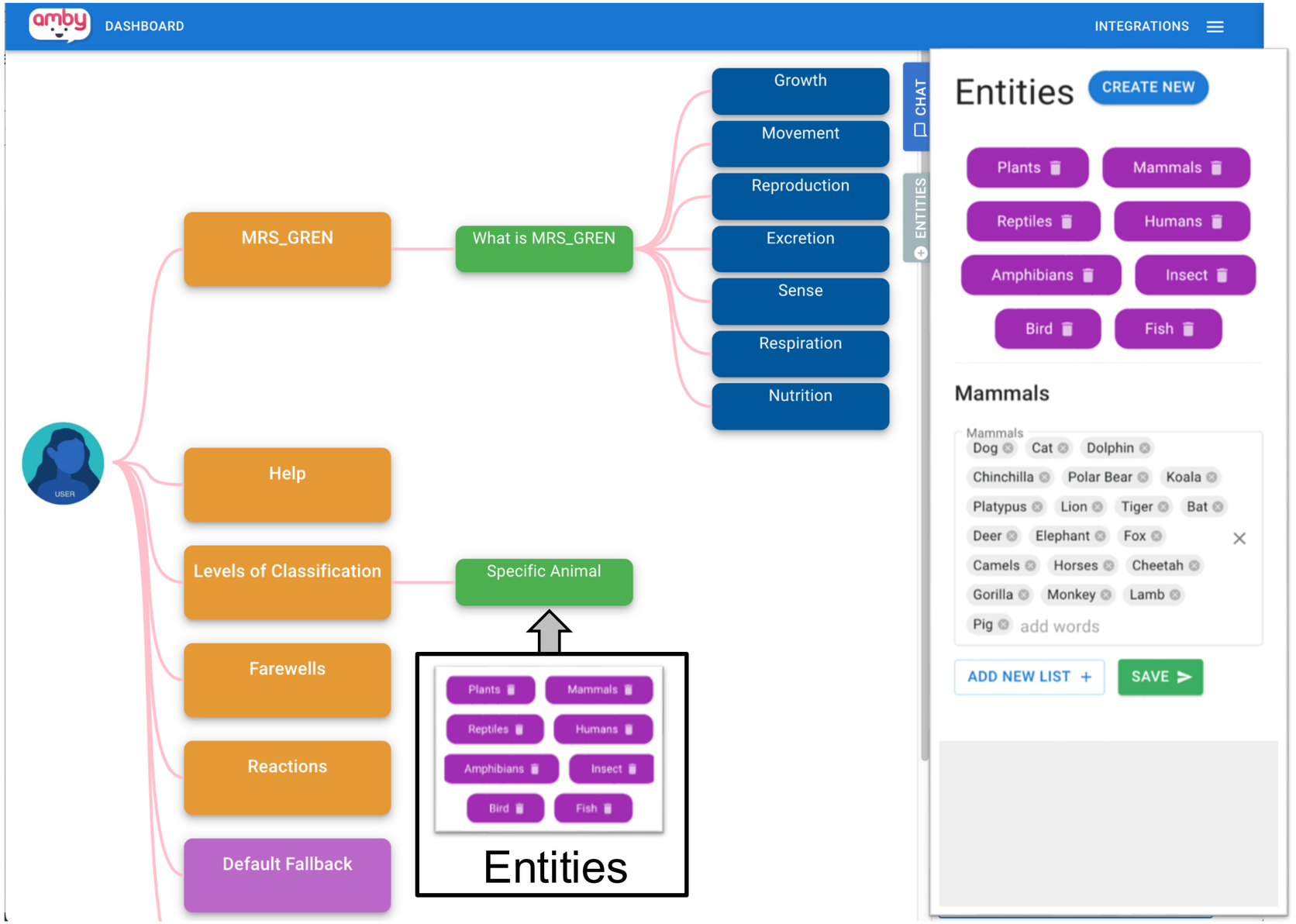
Entity Usage	Did not use (n=18)		Used (n=9)		P original	P adjusted
	Mean	SD	Mean	SD		
Project Ideation	3.11	0.32	2.89	0.33	0.22	0.35
Conv Design	3.25	0.39	3.15	0.27	0.31	0.35
AI Development	3.19	0.25	3.00	0	0.037	0.15
EUS	3.42	0.90	3.23	0.87	0.35	0.35

- P-values were derived from the Mann-Whitney U test comparing the two groups
- Adjusted P-values were using the Benjamini-Hochberg correction

Example 1: *ExperimentBot*

- High average project score: 3.22/4
- High end-user satisfaction score: 4.39/5
- Students noted: efficient and helped better manage their time and effort

Example 2: *LivingThingsBot*



Low project score: 2.89 / 4
Low EUS score: 1.94 / 5

- Lacked information to direct users
- No conversational hints

Results: Entity Feature Perception

- Usefulness
  - Generally useful: save time, efficient, flexible
  - Many considered as a “plus” feature
- Challenges
  - Project topics did not require entities
    - *“Since we were only needing four constellations, it wasn’t broad enough for entities.”*
  - Feature was complexed
    - *“I don’t get how to use entities. Sometimes when I click the dollar sign it immediately goes to the name of my **entity.original**, and I don’t know how to change that.”*

Post-hoc Analysis

- Data aggregated from both conditions
- Outcomes of the classroom intervention
  - Attitudes toward AI
  - Situational interest
  - AI knowledge assessment

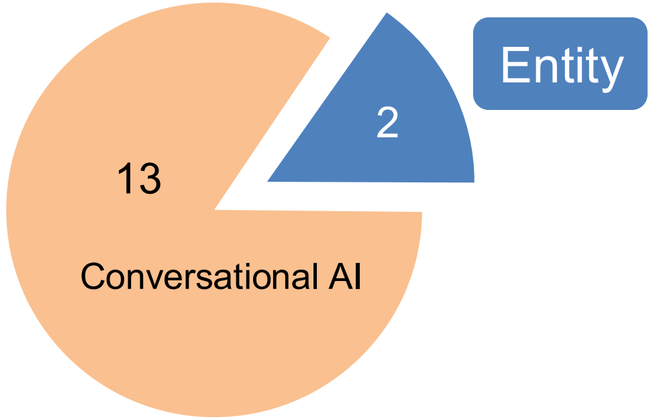
Attitudes and Interest

	Pre		Post		Difference (post-pre)	p value	Effect size
	Mean	SD	Mean	SD			
Ability Beliefs	2.76	0.60	3.18	0.53	0.43 ↑	< 0.0001	0.71
Identity	2.69	0.59	2.60	0.78	-0.09	0.203	0.14
Persistence	2.89	0.55	2.69	0.68	-0.20 ↓	< 0.0001	0.43
Triggered Interest			3.26	0.54	–	–	–
Maintained Interest			3.07	0.70	–	–	–

- P value obtained from paired samples t-test between pre and post responses
- Score scale 1-4
- N=92

AI Knowledge Post-Assessment

- Post-assessment of 15 questions



Student group	Overall AI knowledge (15)		Conversational AI knowledge (13)		Entity knowledge (2)	
	Mean	SD	Mean	SD	Mean	SD
All students (n=98)	0.90	0.07	0.95	0.06	0.56	0.35

Perceived Impact on Science Learning

- Thematic analysis of student written responses in post-questionnaire
- Prompt: Did the conversational AI lessons and activities help you understand science concepts you learn from class? If so, how?

52% Yes	**48% No**
*"It helped my by **refreshing my memory.** It also helped me think about the topic more deeply since I had to **rewrite it in my own words**."*	*"The AI(…) **doesn't tie** into weather, climate, or any science concepts we have learned."*
*"Yes!!! I had to think about **how to explain** something that we learned and **think about different questions**. And you have to understand something to explain it through AI."*	*"Not really because our topic was about astronomy which is space. **In class, we didn't learn much about astronomy**. But I learned somethings like blackholes."*

Discussion

- Entity feature impact
  - AMBY's impact could have overshadowed benefits of the entity feature
  - Less motivated to adopt this "advanced" feature in the classroom
- Challenges identified in formal classrooms
- Design recommendations
  - Incorporating system built-in support
  - Integrating automated assessment and feedback
- Classroom findings are aligned with the educational theories of constructionism and authenticity

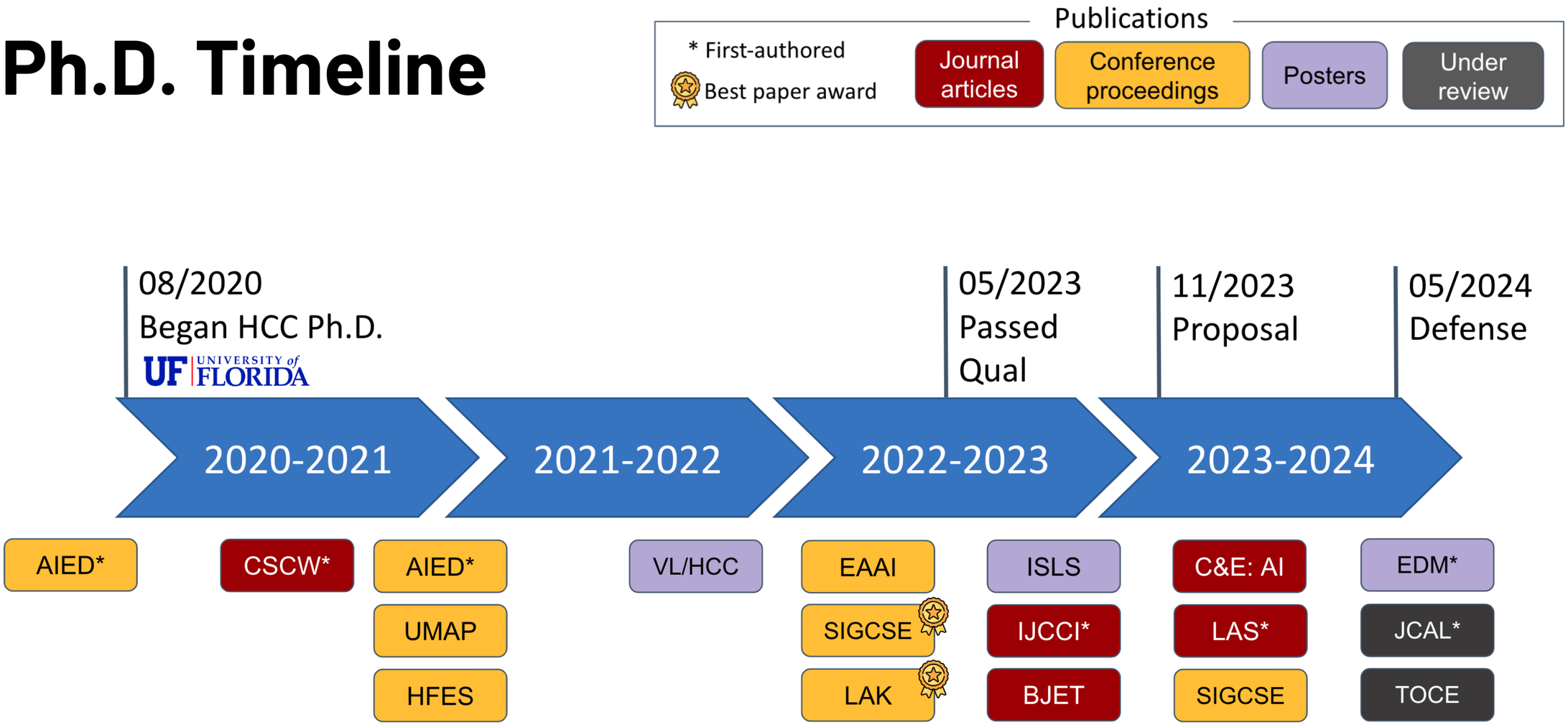
Summary of Contributions

- Design contributions
  - Iterative, children-centered design
  - Design recommendations for AI-authoring tools that support AI education
- Educational contributions
  - A novel approach to teach abstraction to children in the context of AI
  - Integrating AI into middle school science classrooms

Future Work

- Explore more diverse methods of teaching abstraction in the context of AI
  - E.g., Image recognition, generative AI
- Develop AI-based personalized support in the classroom setting
  - Need to approach cautiously
  - Safe and positive learning experiences, mitigate biases
- Scaling up the AMBY program and its curriculum
  - Teacher professional development
  - Integrate AMBY into other subjects

Ph.D. Timeline



Acknowledgements

Advisor

Kristy Boyer

Committee Members

Eric Ragan

Jaime Ruiz

Maya Israel

Project DIALOGS Team Members

DRL-2048480


Kristy Boyer


Maya Israel


Tom McKlin


Joanne Barrett


Mehmet Celepkolu


Timothy Brown


Lydia Pezzullo


Gloria Katuka


Yukyeong Song


Amit Kumar


Sunny Dhama


John Tran Hoang


Christine Fry Wise


Carly Solomon

Special thanks for helping with Classroom study:


Priya Prasad


Shan Zhang


Toni Earle-Randell


Maedeh Agharazi


Wesly Menard

Current and Past LearnDialogue Members


Timothy Brown


Julianna Ruiz


Mehmet Celepkolu


Maedeh Agharazi


Lincoln Lu


Joseph Wiggins


Lydia Pezzullo


Yingbo Ma


Gloria Katuka


Amanda Griffith


Amogh Mannekote


Toni Earle-Randell


Chandler Wiggins


Alex Johnson


Wesly Menard


Mollie Brewer


Sunny Dhama


Abraham Banos Lombardero


Carly Solomon


Amit Kumar

Family and Friends

UF Herbert Wertheim College of Engineering
UNIVERSITY of FLORIDA

Q & A

Xiaoyi Tian

Ph.D. Proposal Defense

tianx@ufl.edu