

# HRI Reading Group

@ Instituto Superior Técnico  
Spring 2019

Meeting #15 (July 12, 2019)

# Paper

Tom Hitron, Yoav Orlev, Iddo Wald, Ariel Shamir, Hadas Erel, and Oren Zuckerman.  
2019. **Can Children Understand Machine Learning Concepts?: The Effect of Uncovering Black Boxes.** In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). ACM, New York, NY, USA, Paper 415, 11 pages. DOI: <https://doi.org/10.1145/3290605.3300645>

# Background

- **Children constantly learn from experience** by interacting with the physical world around them.
- This direct exploration contributes to the **construction of mental models**, which are **conceptual and operational representations of phenomena** and processes in the world.

# Problem

- **Direct exploration is limited when processes are "black-boxed"** (i.e. hidden from the user), making it harder to construct accurate mental models.
- **Uncovering too many processes can interfere with the learning process** as a novice learner may be overwhelmed

# Supervised ML and Classification Problem

- Relatively less complex and common in real world applications

Supervised ML can be defined as a pipeline consisting of four building blocks:

- Data Labeling
- Feature Extraction
- Model Selection and Validation
- Evaluation

# Data Labeling

- Sample Size (sufficient amount of data);
- Sample Versatility (sufficient variety of positive examples within a class);
- Negative Examples (inclusion of examples that do not belong to a class)

# Feature Extraction

- Preprocessing the data in order to simplify classification and speed up computation
- Finding the useful features that are fast to compute rather than feeding the raw data to the algorithm

# Model Selection and Validation

- The consideration and testing of different model types (and their parameters) in order to find the best one for a particular application

# Evaluation

- Testing the trained ML model with new data to evaluate the quality of the chosen model

# What does it mean “Uncovering Black Boxes”?

In some model what you have is an implicit representation of the model

There is a difference between interpretability of the process and interpretability of the output.

They are different concepts but both are important.

In this paper they talk about the way with interact with technology: the states and the mechanisms of the algorithm.



# What are the associated **benefits** and **disadvantages** of a deep understanding about technologies from an early age?

## Benefits

We all have theories about how things work, having this systems will allow children to have a better/clear understanding of the system

Children are not only consumers of technology but become designers from an early age.

## Disadvantages

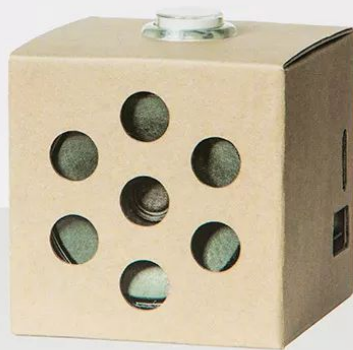
They become nerds



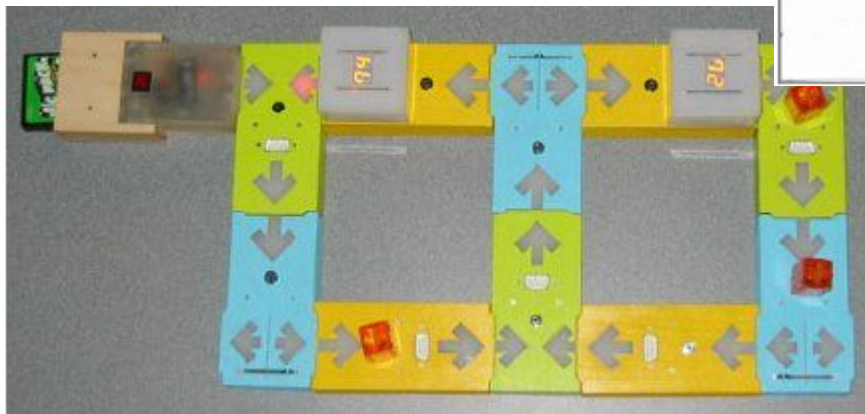
Made by you  
with Google

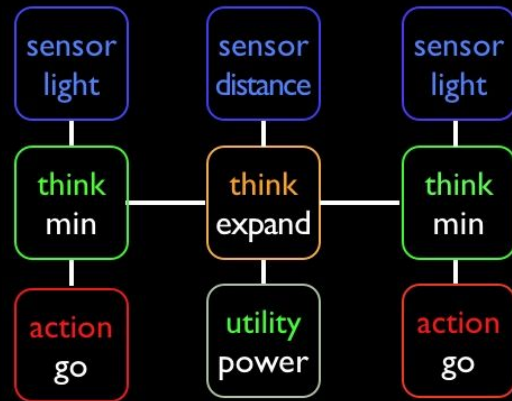
# Voice Kit

Do-it-yourself intelligent speaker



Raspberry Pi  
Included





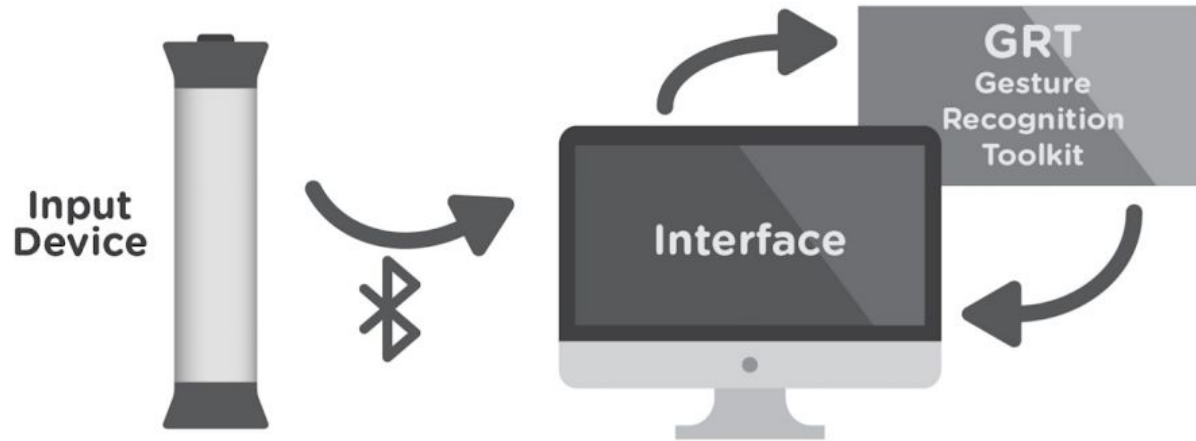



# Preliminary Work

**Wizard-of-Oz experience** designed to give children feedback when "training" a device:

- **Lesson Learnt:** direct experience with accessible ML building blocks have the potential to enhance children's understanding of basic ML concepts

# Gest, a ML gesture recognition system





digitally-enhanced devices for outdoor play

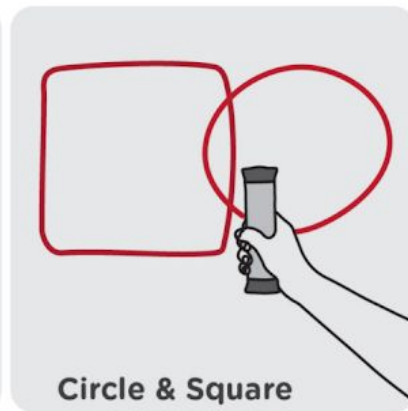
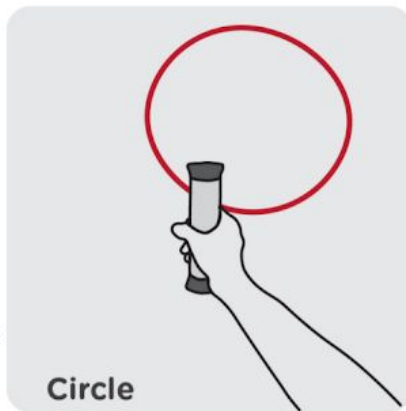
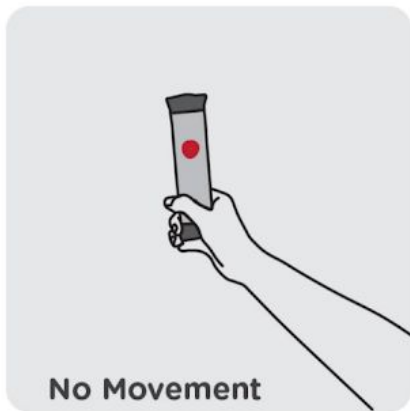
Scratch nodes



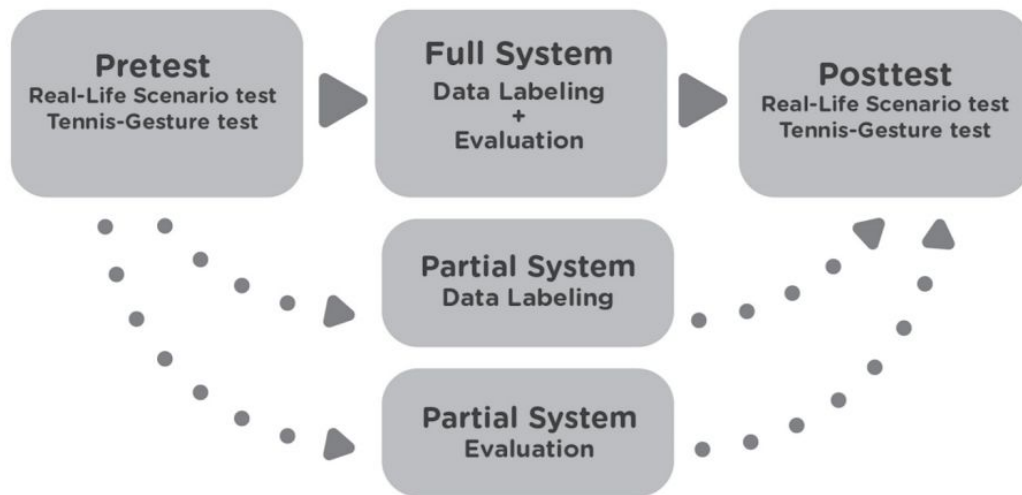
# Design Principles

- Design Principle 1: Low Floor
- Design Principle 2: Uncovering black-boxes
- Design Principle 3: Promote Iterations
- Design Principle 4: Promote Self-generated Knowledge

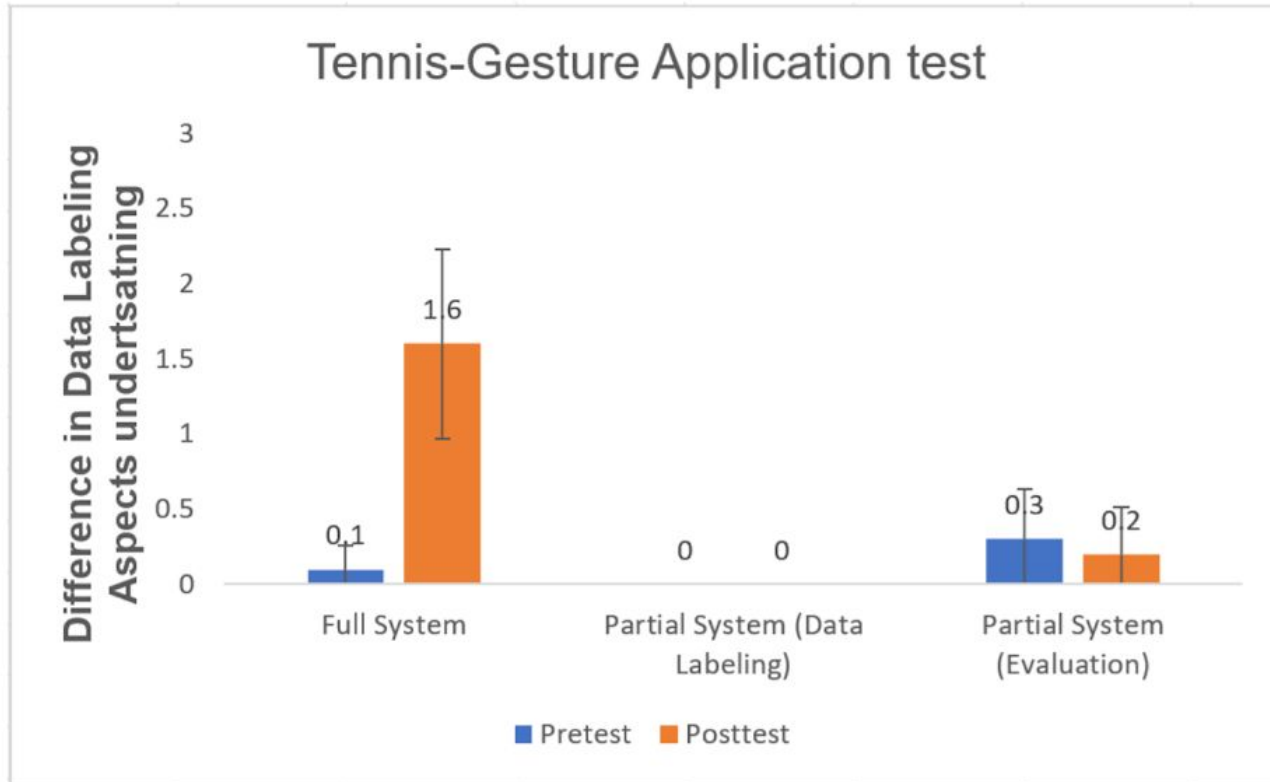
# Learning Task



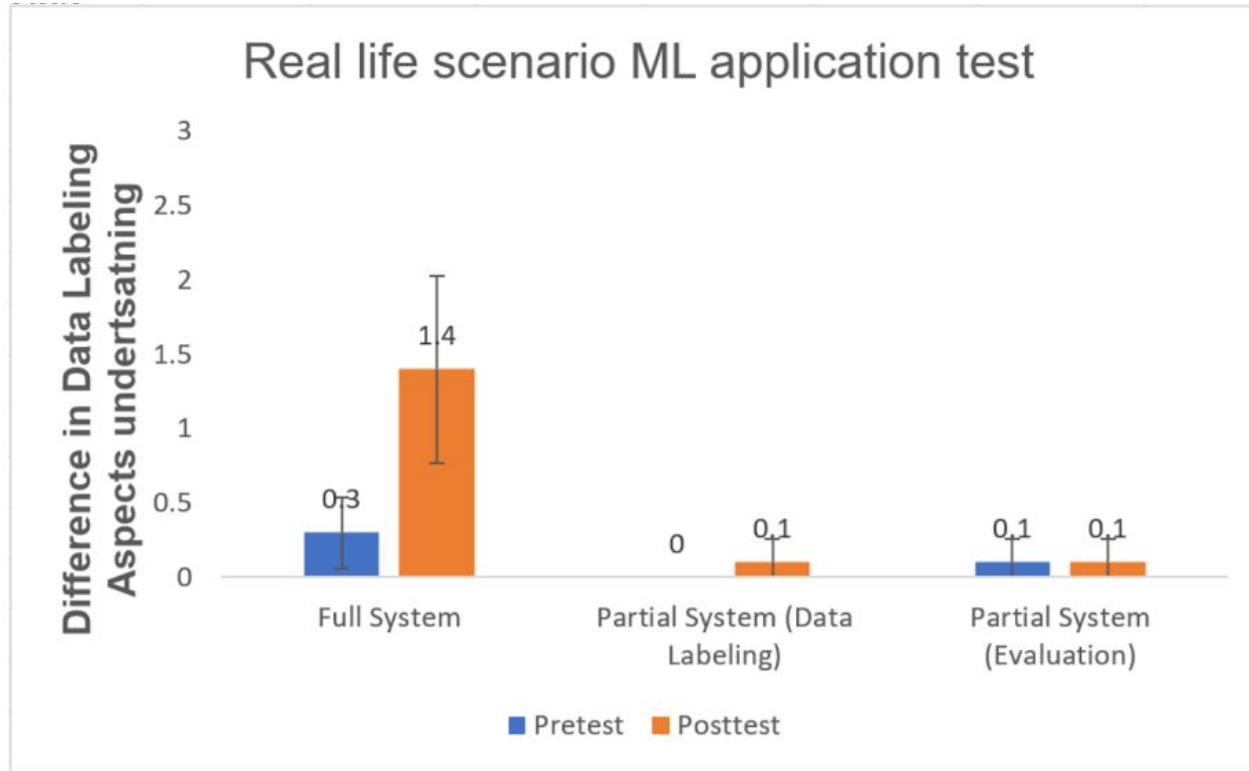
# Research Design and Method



# Results



# Results



# Results

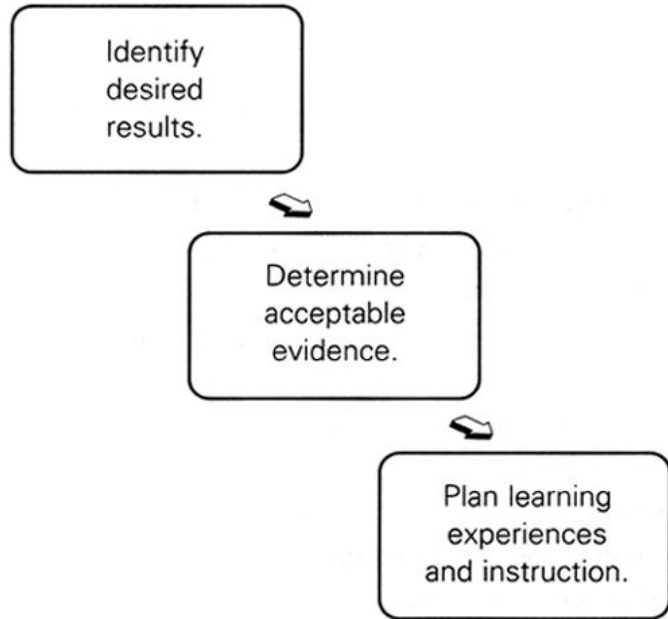
**Table 1: Number of children who stated each data labeling aspect in their explanation**

		Real life scenario ML application test		Tennis-gesture application test	
		pretest	posttest	pretest	posttest
Sample Size	Full System	3	6	1	7
	Partial System: Data Labeling	0	1	0	0
	Partial System: Evaluation	1	1	2	1
Sample Versatility	Full system	0	5	0	6
	Partial System: Data Labeling	0	0	0	0
	Partial System: Evaluation	0	0	1	1
Negative Examples	Full system	0	3	0	3
	Partial System: Data Labeling	0	0	0	0
	Partial System: Evaluation	0	0	0	0

They provided no feedback on some conditions which can be a limitation in the differences between conditions

Activity:

## Design a learning activity for children that involves system transparency



### Setting & Outside constraints

- Timing and Logistics
- Students
- Situational Constraints
- Schedule

### Learning Objectives

- Which kind of task?
- What are the design principles?
- How to measure the learning task?

### Role of Transparency?

## Group 1

Teach children how important the sample size and the variability of the sample is so that algorithms do not discriminate people.

Child trains the robot by showing pics of people that the robot needs to classify. Show a variety of pics from people, g.e.g, different cultures and nationalities

Measurements: accuracy of the robot's classifier

We are teaching the child how to build a training dataset

For example, give the child 100 pictures and has to select 10

We compare pre and post robot interaction to see if the dataset they select is more balanced in the post

## Group 2

Problem: children need to understand the dog's needs

Solution: use a game to express what the dog wants

E.g., if the dog is waving its tail, the game will provide different options to what a dog could need (going for a walk, the owner arrived, wants to be pet). The child needs to try one of the options provided by the dog.

Measurement: score of the game translates how well they are treating the dog.



# HRI Reading Group

@ Instituto Superior Técnico

Meetings will resume in the Fall Semestre