Object Individuation and Physical Reasoning in Infancy: An Integrative Account

Authors: Renée Baillargeon , Maayan Stavans , Di Wu , Yael Gertner , Peipei Setoh ,
Audrey K. Kittredge & Amélie Bernard

Presented by: Priya Thanneermalai

Motivation

- Previous research in this area involved versions of a typical experiment with a screen and 2 objects
- It was not until the end of the infant's first year, that they realized that an object had gone missing
- Any modifications to this experiment gave inconsistent results
- This paper tries to explain this situation using Physical Reasoning to explain Object Individuation
- Helps in understanding how infants represent object information across events

Physical Reasoning (PR)

- Infants are born with PR, an abstract computational system with a skeletal causal framework for reasoning and learning about physical interactions of objects
- Infants are not aware that they have this!
- Persistence: Objects persist in time and space

How Do Infants Learn About Events?

- PR system builds a specialized physical representation of the event.
- More detailed as infants grow

Event Categories

-Occlusion : Event in which an object occludes another object

-Support: Event in which an object supports another object

-Collision: Event in which an object hits another object

-Containment: Event in which a container contains an object (Box, cup)

-Each of the objects in these events are viewed as playing specific roles

Explanation Based Learning (EBL)

-4 steps

- Learning is triggered when infants build similar physical representation of 2 events but with contrastive outcomes

- Infants search for the conditions that map onto these outcomes until they detect a possible condition-outcome regularity

- Infants then construct an explanation for the observed regularity

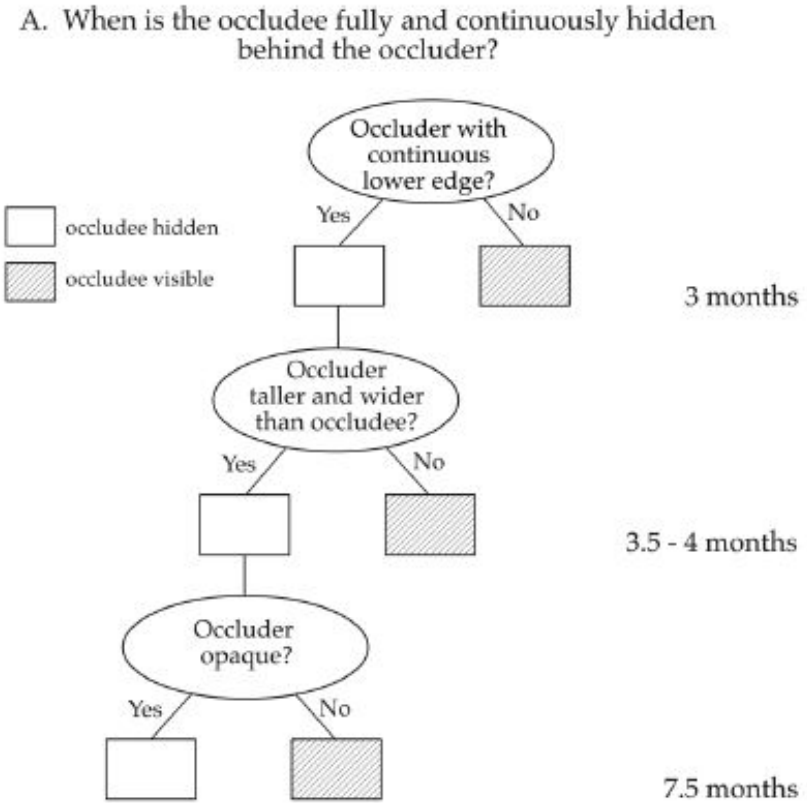
- Finally, the explanation suggests a general causal rule

Variables

- EBL also helps identify variables.

- Infants do not always get to the right conclusion at the start. So, they revise their flawed rule through the identification of relevant variables.

- Example of variable: Lower Edge Discontinuity, Height, Color, Pattern

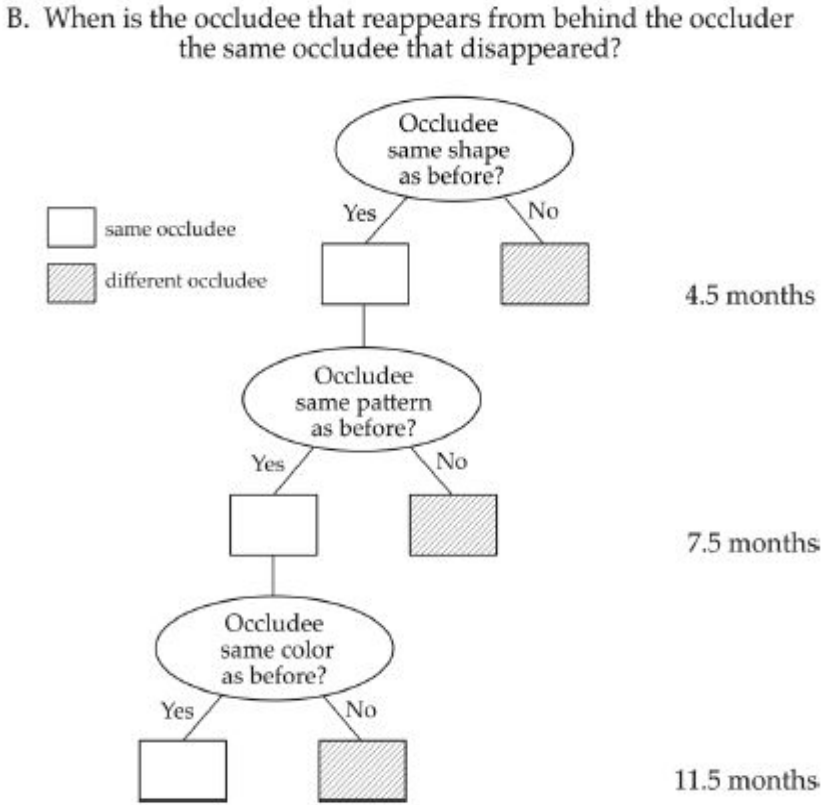


Additional Vectors

- Infants often notice contrastive outcomes having to do with other facets of the events, leading to formation of additional vectors

- When and where will the occludee reappear from behind the occluder?
- Is it the same occludee that disappeared?

- Discussion question 1: Difference between additional vectors and variables?



Decalages

- Variables are not transferred across event categories, even when equally relevant leading to lags or decalages.
- Variable height is identified in occlusion events at 3.5 months, but it is only identified in containment events at 7.5 months and at 12 months in covering events.

Learning What vs Learning How

- Infants know what to add to their internal physical representation but not the how to reason about that information.
- In the first year of life, the PR system involves 2 distinct processes: interpretive and predictive.

How Do Infants Represent a Physical Event?

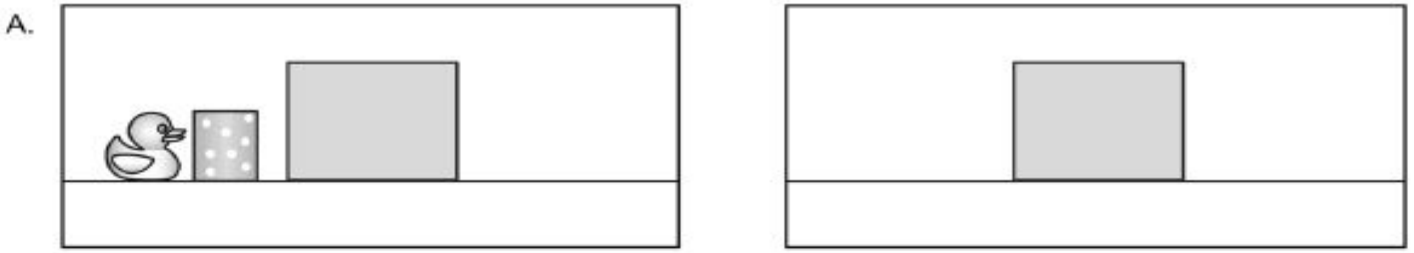
- 2 layers: structural and variable
- Structural layer: includes generic information about any event
- variable layer :includes more specific information that differs across events

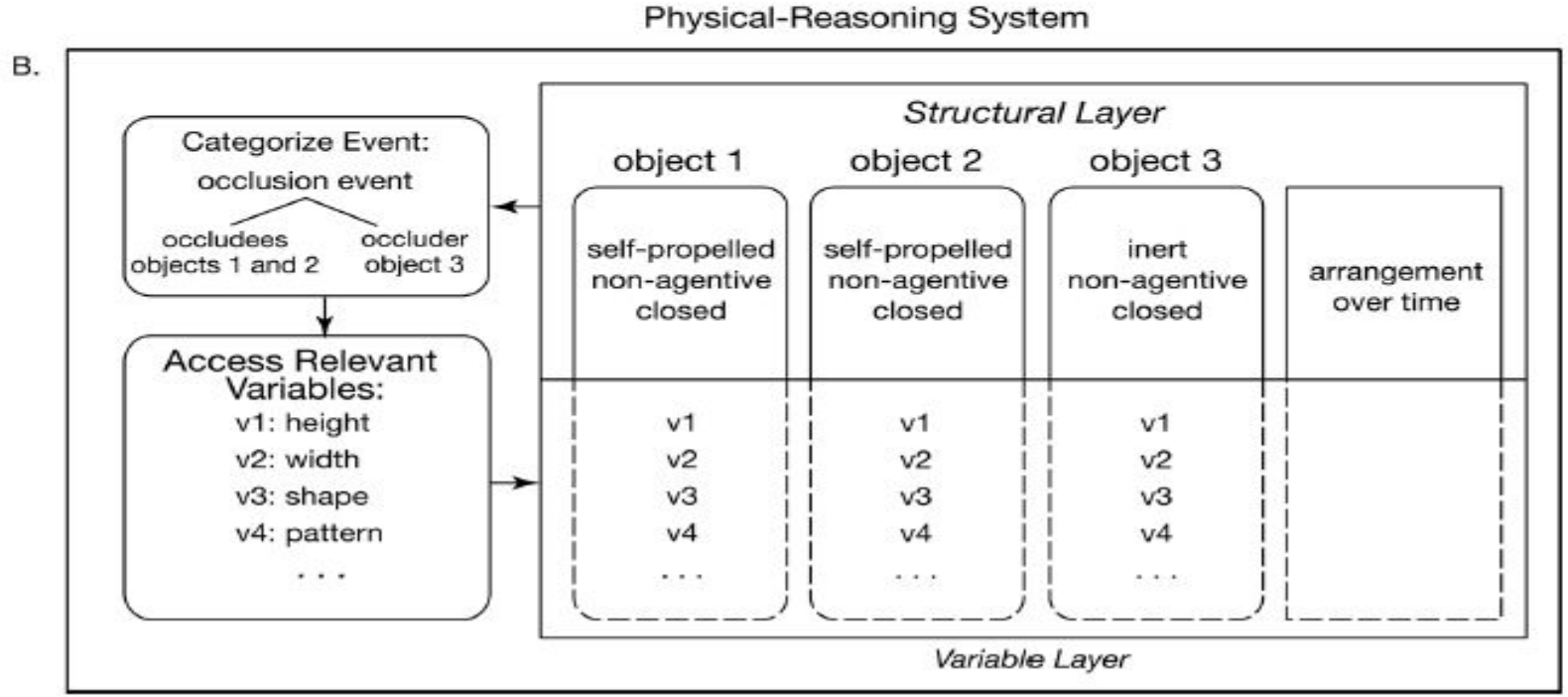
Structural layer

- Includes spatiotemporal (arrangement of the objects in the event) and categorical (what kinds of objects are involved in the event) information
- As event categories are identified, they change the structural layer of the physical representation
- Experiment: 3 objects (yellow toy duck, red block, green screen) are placed in a row. Duck and block move behind the screen.
- Spatiotemporal: 3 objects occupying separate locations.
- Categorical: Closed, self-propelled (move by themselves)
- On movement, event is categorized as occlusion with screen being the occluder and other 2 toys being the occludee.

Variable layer

- Infants identify variables that make the physical representation more richer and detailed.
- When an event occurs and the infant is watching it, the PR system first represents the structural information about the event and uses this information to categorize it.
- Next, the PR system taps the list of variables for that event category that have been identified as relevant for predicting outcomes.
- This variable information is then included in the physical representation.


A.


Physical-Reasoning System
B.
Categorize Event:
occlusion event
occludees objects 1 and 2
occluder object 3
Access Relevant Variables:
v1: height
v2: width
v3: shape
v4: pattern
. . .
Structural Layer
object 1
object 2
object 3
self-propelled non-agentive closed
self-propelled non-agentive closed
inert non-agentive closed
arrangement over time
v1
v2
v3
v4
. . .
Variable Layer

How Do Infants Represent a Sequence of Two Physical Events?

- What happens when an object is first involved in one event category and then another sequentially? Are there entirely different PR systems for each?
- Variable information from the first event is carried by the PR system into the second event.
- No carryover for different objects in sequential events.
- Positive and negative carryover effects
- Binding process: Structural and variable layers are bound tightly and carried from one event to the next. Eg. Screen is fully removed after the occlusion event

How Do Infants Reason About Physical Events?

- This comes after physical representation, how do they explain events?
- Eg. After the occlusion event, what if the screen is lifted and only 1 object is seen?
- Eg. What if it is a different object, when the screen is lifted to reveal the objects?
- Research indicates that infants are surprised by such Variable Persistence Violations
- Only if infants have identified that variable is relevant to that event category in their physical representation

- After removing the screen, if 3 objects appeared for eg. the duck, the block and a toy car, infants concluded that the toy car was already hidden behind the screen before lifting it.
- If there was a second screen nearby, and one of the objects disappeared upon lifting the first screen, they assumed that it was hidden behind the second screen.

Findings of Xu and Colleagues: 10 month olds

- Experiment 1: In each of the test trials, each object emerged from behind the screen and returned behind it multiple times.
- Baseline Trial 1: Remove screen to reveal 2 objects
- Baseline Trial 2: Remove screen to reveal 1 object
- Test Trial 1: Remove screen to reveal a bottle and a ball.
- Test Trial 2: Remove screen to reveal a cup and a book.
- Infants looked longer every time 2 objects appeared after the screen was lifted in all the trials.

- Experiment 2: Same as previous experiment but only a single object was involved every time.
- When one object appeared instead of 2, after the screen was lifted, the infants readily generated that the other was still hidden behind the screen.
- Experiment 3: Same as the previous experiment but both objects were brought out together at the start meaning clearly that there were 2 identical objects
- Infants actually looked longer at 2 objects being revealed than one object which is surprising
- All in all, When two distinct objects emerge in alternation from behind a screen, 10-month-olds have no clear expectation as to whether one or two objects should be revealed when the screen is removed.

Findings of Xu and Colleagues: 12 month olds

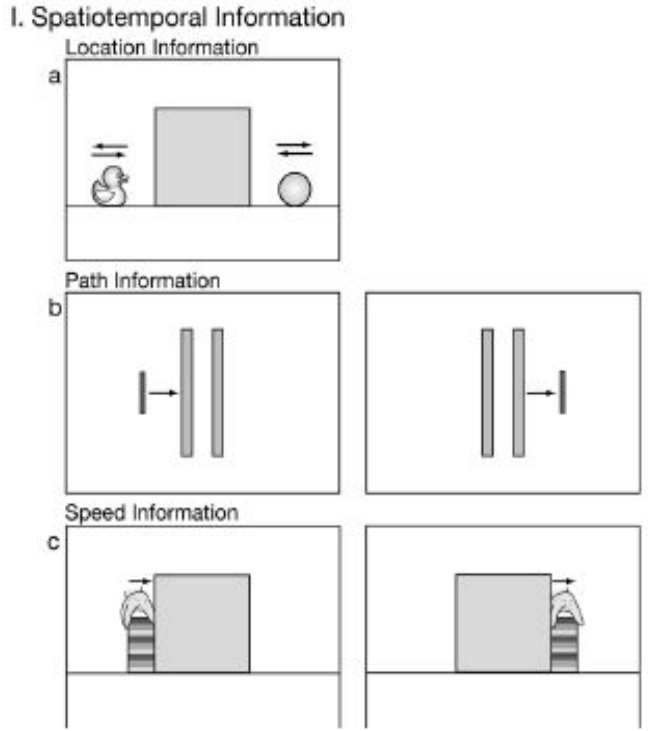
- They looked longer at 2 objects in baseline trials (surprise) than one object but looked equally in both test trials.
- This happened only when the two objects belonged to different object categories.
- Both these experiment sets are inconsistent to what the paper explains so far.
- But if infants believe that two objects are present behind the screen, why do they not expect to see two objects when the screen is removed in baseline trials? Why are infants not surprised when a single object is revealed (Why do they assume the other is hidden)?

Assumptions

1. The Carryover of Object Information Breaks Down When Inconsistencies Exist Between the Structural and Variable Layers of an Event's Physical Representation
2. By the End of the First Year, Categorical Information in the Structural Layer Begins to Include Taxonomic Descriptors
3. Reconciling an Inconsistency Between the Structural and Variable Layers of an Event's Physical Representation Requires Creating Distinct Spatiotemporal (where) Descriptors

Infants' Use of Spatiotemporal and Categorical Information

- Types of Spatiotemporal Information
- Location: 2 different objects appear from behind screen, so expect to see 2 objects
- Path: Object disappears behind a screen and reappears behind another screen, expect to see 2 different objects
- Speed: Object disappears behind left edge of the screen and immediately reappears on the right edge, expect to see 2 objects

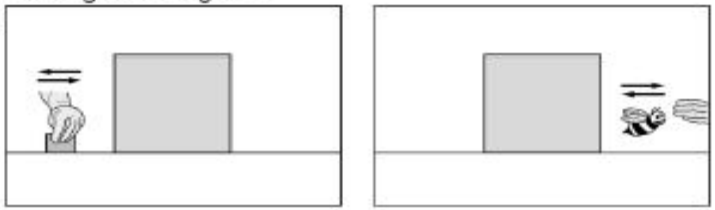


- Categorical Information : 3 types of Categories
- Ontological Category: block (inert) appears from left edge first, then bee (self propelled) appears from right edge of screen; expect to see 2 objects
- Functional Category: Closed like balls or blocks, internal openings like rings and tubes.
- Taxonomic Category: Labels or names of objects

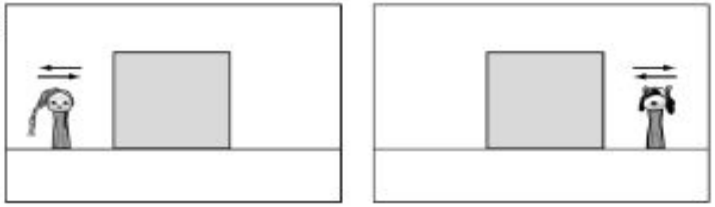
II. Categorical Information

Ontological Categories

d

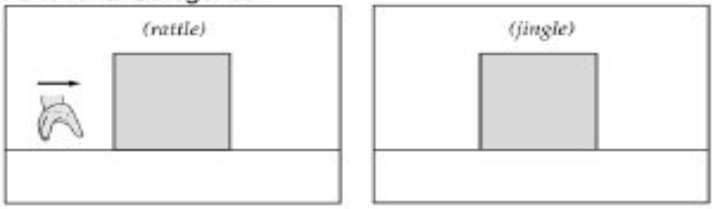


e

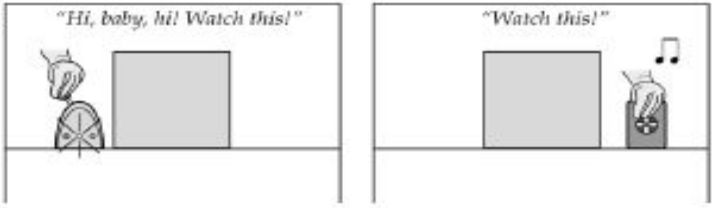


Functional Categories

f

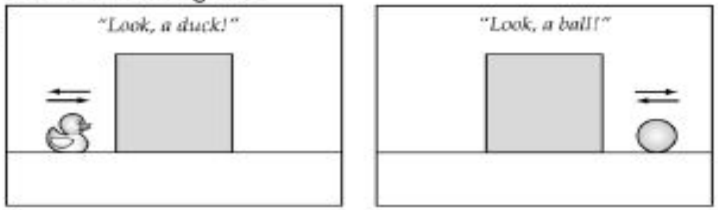


g

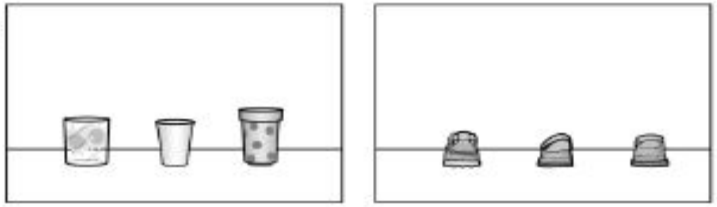


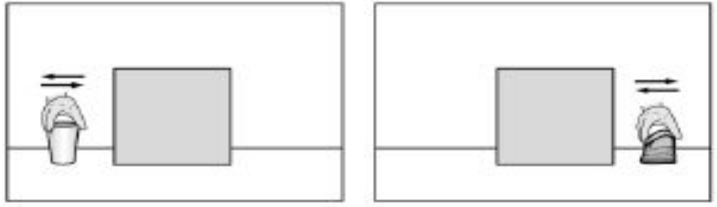
Taxonomic Categories

h



i





Do Infants Realize There Are 2 Objects During the Event

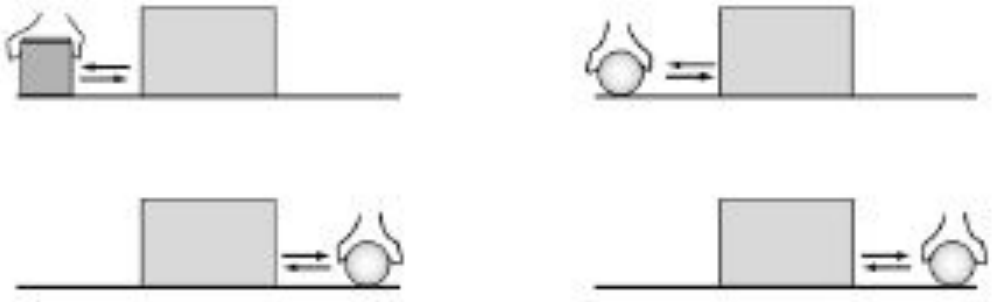
- Possibility 1: No clear expectation/ confusion about how many objects are behind the screen during occlusion
- Possibility 2: Inconsistencies matter only when info needs to be carried over to next event. And only bundled up info is carried over which does not happen till the next event has started.
- Experiments with 2 screen tasks give us the answer (the latter possibility)

Two-Screen Tasks With Opaque and Transparent Screen

- A box and a ball emerge from behind the screen and then the opaque screen is removed, revealing a transparent screen behind it and just one object
- Same experiment was done with 1 ball emerging from behind the screen.
- Result: Surprise in the box ball condition but not in the ball ball condition
- Discussion question 2: Why?

A. Box-ball Condition Ball-ball Condition

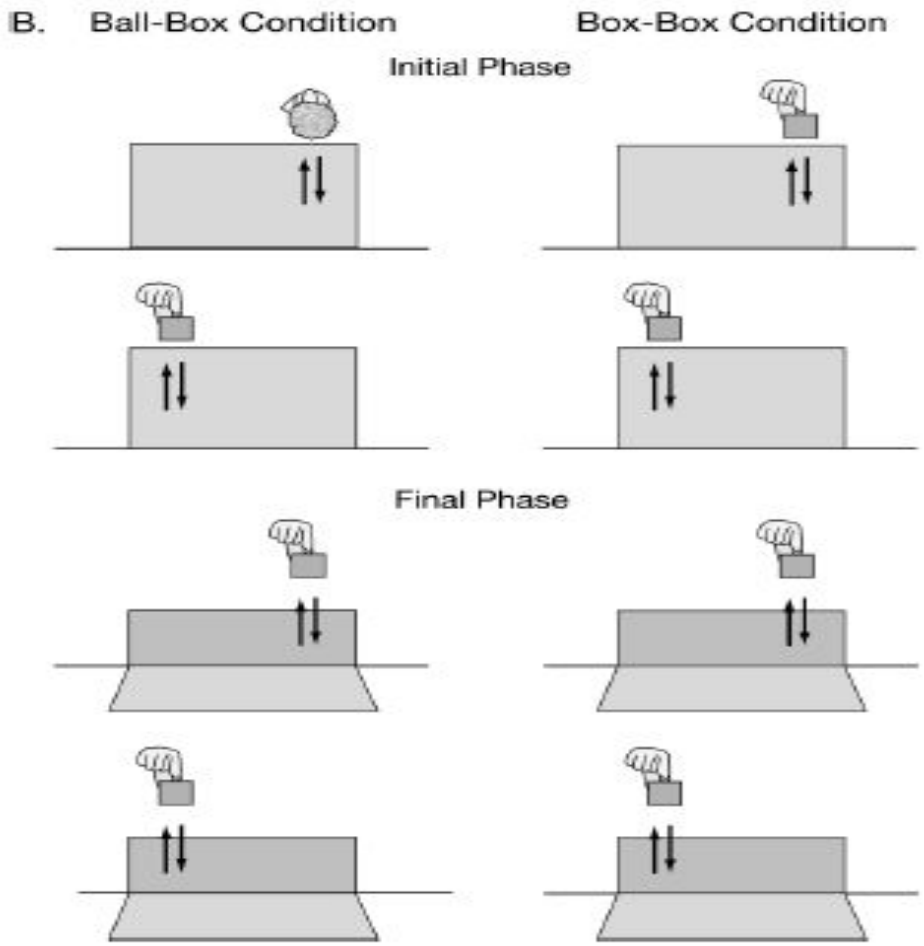
Initial Phase



Final Phase

Two-screen Tasks With Two Opaque Screens

- A ball emerged from the top right corner and a box emerged from the top left corner of the screen. The screen was removed to reveal another completely different screen and then a box emerged from both corners.
- Same experiment but with the same box.
- Result: Surprise in the ball box condition than the same box condition
- Discussion question 3: Why?
- Follow up: Instead of the second screen, a container was revealed and only the box is lifted out
- Discussion question 4: What do you think is going to happen now?


B.
Ball-Box Condition
Box-Box Condition
Initial Phase
Final Phase

Piazza and Class Discussion summary

- The experiments need to be conducted in a very controlled way so that we do not misread the reactions of infants, given that they cannot communicate clearly verbally.
- A few people agreed that proving the correlation of staring at something longer indicating an expectation being unfulfilled is tough to prove.
- However, in class we talked about how this method of experimentation in infants is indeed proven to work. Also, this paper mitigates the effect by letting the infants explore all new objects used in the experiments so that they do not find it novel and are not fascinated by it.
- Some people pointed out how in the previous papers, researchers used time as a variable to measure affordances

- There was some confusion which was clarified by later comments about why 10 month olds could not see 2 objects as 2 when they were of different categories and colors while 12 month olds could. The explanation is that 10 month olds do not identity color or certain categories and thus see 2 objects as 1.
- There was analogy between affordances and object individuation being discussed, where some people thought that replicating it into a machine model would be really hard and there would be uncertainty about how to start with original knowledge.
- A few people wondered if infants know what can they do with the objects. Also, in terms of applications of this paper's concept, a 3D tracking algorithm was referenced where objects are associated with themselves in a different frame with a feature identity network.

- Object path during occlusion is predicted using Kalman filter and reassociated to the original object after occlusion. Physical reasoning and knowledge of object properties and behavior also play an important role.
- In computer vision, object category recognition and classification are important tasks for many applications, and have been the focus of extensive research.
- In terms of variables, attention and attention shifting are noteworthy. In order to track and reason about object states, individuals need to selectively attend to relevant objects and features, and to shift attention between objects as needed.
- In computer vision, attention-based approaches have been proposed to improve object tracking and detection, by selectively attending to the most informative parts of the image (Wang et al., 2020).

- A couple of videos were discussed and one of them was played in class.

  How Infants Learn: Dr. Renee Baillargeon (https://www.youtube.com/watch?v=Zd7OIDm_btM)

  Understanding Object Permanence

  (https://www.youtube.com/watch?v=b1tQOR5L0iI)

- The ability to recognize and track objects in a video stream is a critical component of many CV applications. CV algorithms must be able to distinguish between objects in a scene and track their movements over time, even when there are occlusions or other disturbances. Layers discussion relates to multi-layer neural networks to extract both low-level and high-level features from image and video data.

Findings from the 2 Screen Experiments

- Only after the occlusion event, they do not hold clear expectations of how many objects are present when they are shown 2 objects but only 1 is revealed
- However, during the event, they have information in the variable layer and thus correctly expect to see 2 objects
- If they are 2 identical objects, infants fail to detect a violation when only 1 object is revealed

Predictions

- 10 month olds that fail to detect a violation when only 1 object is revealed (in the experiment with the duck and ball) will also fail to detect a violation when no object is revealed (Xu's finding, assume that the other object is still hidden)
- Similarly, 12 month olds when tested with a small and a big ball fail to detect violation when 1 object is revealed will fail to detect a violation when no object is revealed.
- Whether they succeed or fail depends on: PR system identifying the necessary variables and inclusion of necessary information in the variable layer

Key Takeaways

- When 2 distinct objects emerge from behind the screen, infants realize that there are 2 objects during the event as long as objects belong to different categories
- Object information can only be carried over if it is internally consistent
- After screen is removed, infants expect to see 2 objects if during occlusion, spatiotemporal and/or categorical info states 2 objects are present or categorical info points to 1 object but variable layer points to 2 objects
- By age 1 year, infants encode taxonomic categories
- After screen is removed, infants expect to see 2 objects when structural layer includes 2 distinct categorical (what) or spatiotemporal (where) descriptors