Online data collection for developmental research

Kim Scott | Thesis defense | September 28, 2017

Committee: Laura Schulz | Rebecca Saxe | Liz Spelke | Josh Tenebaum

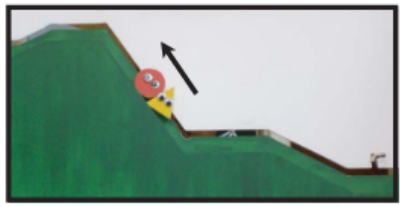
Online data collection for developmental research

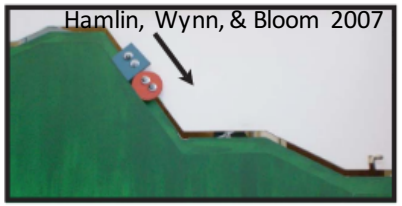
Kim Scott | Thesis defense | September 28, 2017

Committee: Laura Schulz | Rebecca Saxe | Liz Spelke | Josh Tenebaum

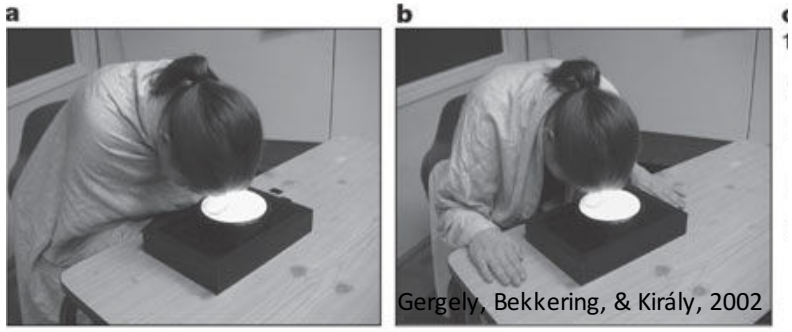
Practical challenges for developmental research: posing questions so babies can answer


Werker, Polka, Pegg, 1997

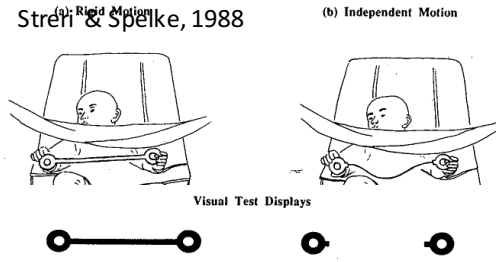



Hamlin, Wynn, & Bloom 2007


Meltzoff & Borton, 1979

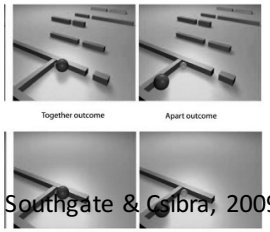


Gergely, Bekkering, & Király, 2002

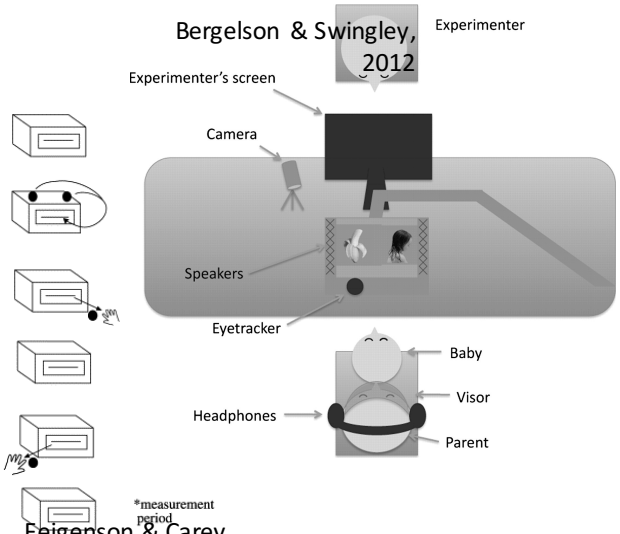


Streri & Spelke, 1988

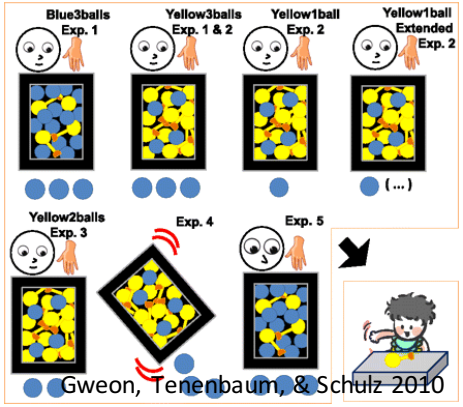

Warneken & Tomasello



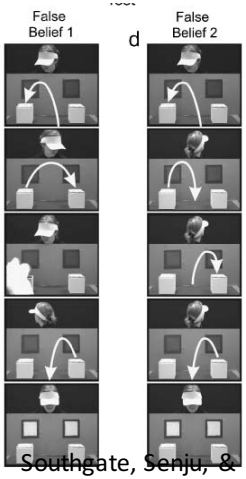
Southgate & Csibra, 2009



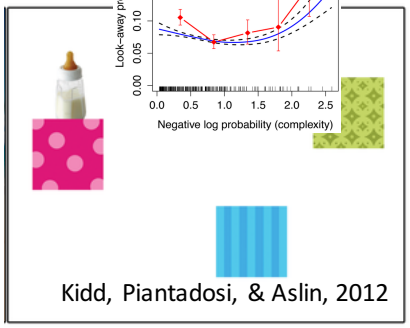
Bergelson & Swingley, 2012
Feigenson & Carey, 2005



Gweon, Tenenbaum, & Schulz 2010



Southgate, Senju, & Csibra, 2007



Kidd, Piantadosi, & Aslin, 2012

Practical challenges of developmental research: getting kids into the lab

- Takes a lot of time to get kids into the lab
  - Small sample sizes
  - Large effects, few conditions
- Even harder to bring them back!
- Recruiting special populations locally
- Kids behave differently in the lab
- Kids who come in aren't representative
- Socially-transmitted methods

https://www.mass.gov/how-to/order-a-birth-marriage-or-death-certificate

Access to records of births to unwed parents is restricted to the:

- Child
- Parents
- Father not listed on the record (this requires documentation supporting paternity)
- Legal guardian (this requires documentation supporting guardianship)
- Legal representative (this requires documentation supporting representation)

Image sequence: lab tour, Harvard LDS
https://software.rc.fas.harvard.edu/lds/parents/visit-the-lab/

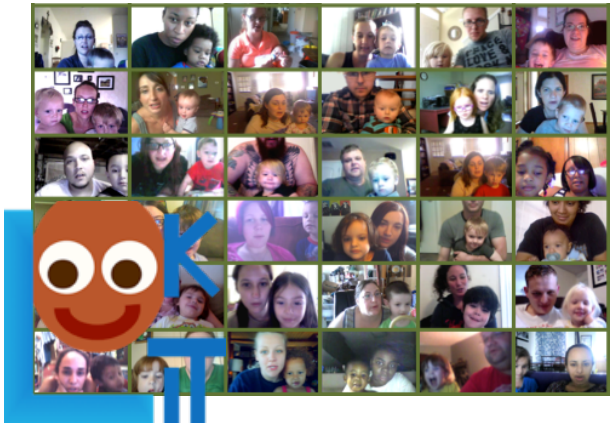
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

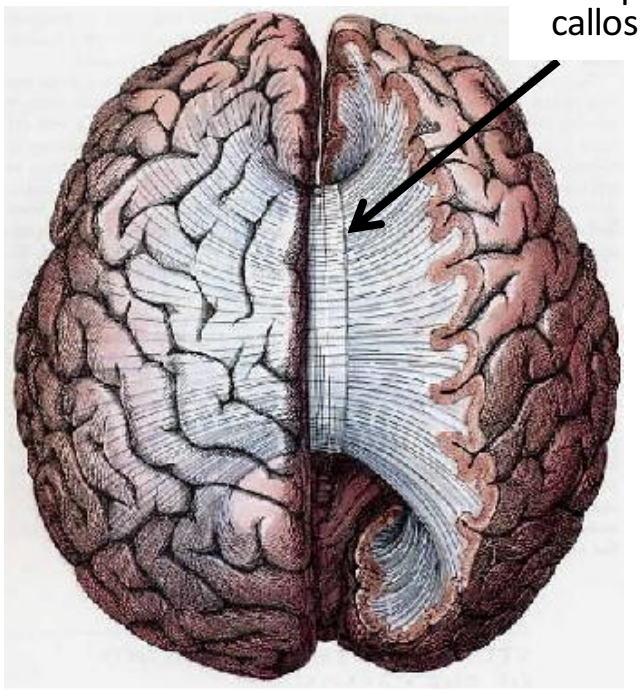
- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



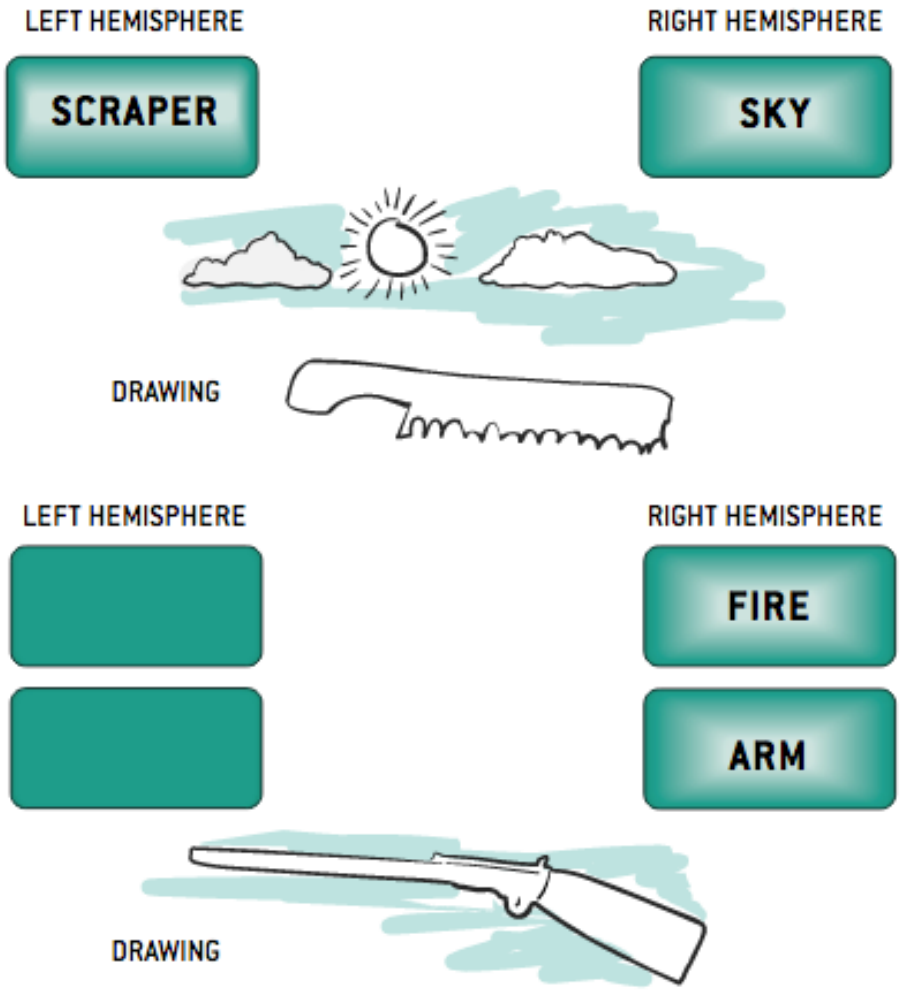
Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Case study: evaluating interhemispheric integration in infancy



http://hubel.med.harvard.edu/book/b34.htm



Gazzaniga, M.S. The Split Brain Revisited. Scientific American, 2002.

Do we start as “little split-brains”?

functional deficits…

Up to age 7:

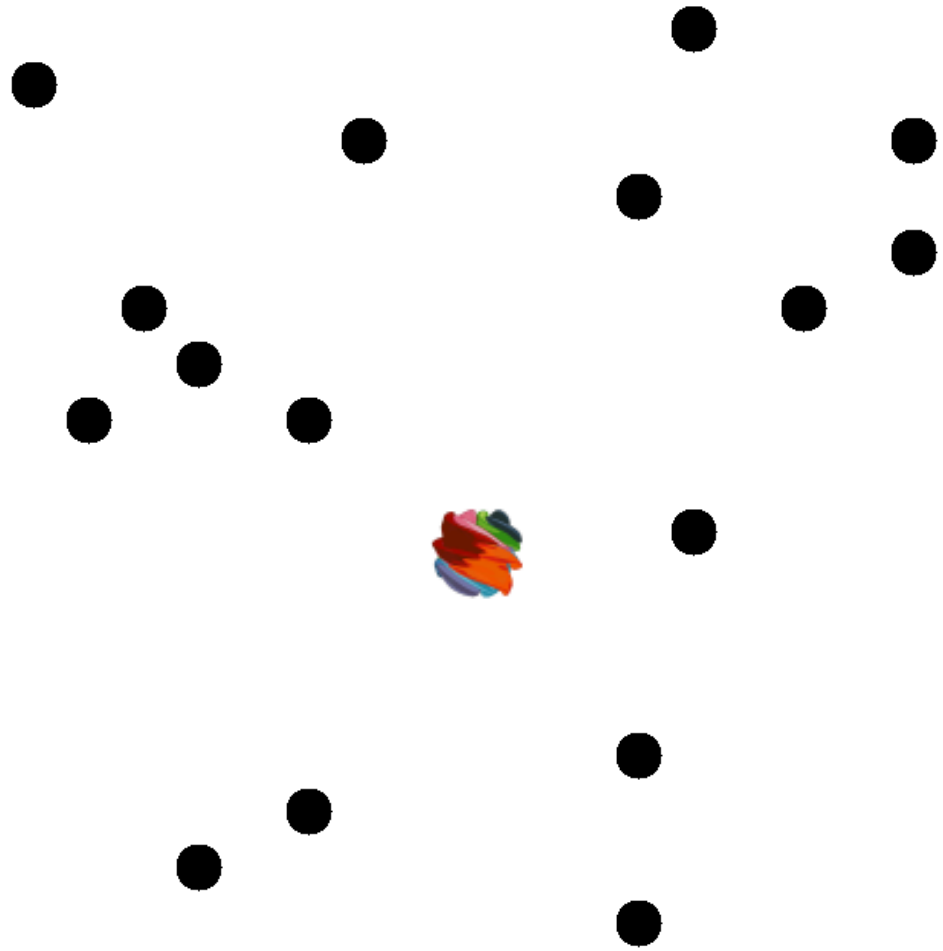
- Fail to transfer a manual task learning in one hemisphere to the other (Chicoine, Proteau, & Lassonde, 1992)

Up to 24 months:

- Fail to learn rules requiring comparison of shapes across the midline (Liegeois and de Schonen 1996; Leigeois et al. 2000)

but what do they experience?

Do infants combine number across the midline?



Why use approximate number?

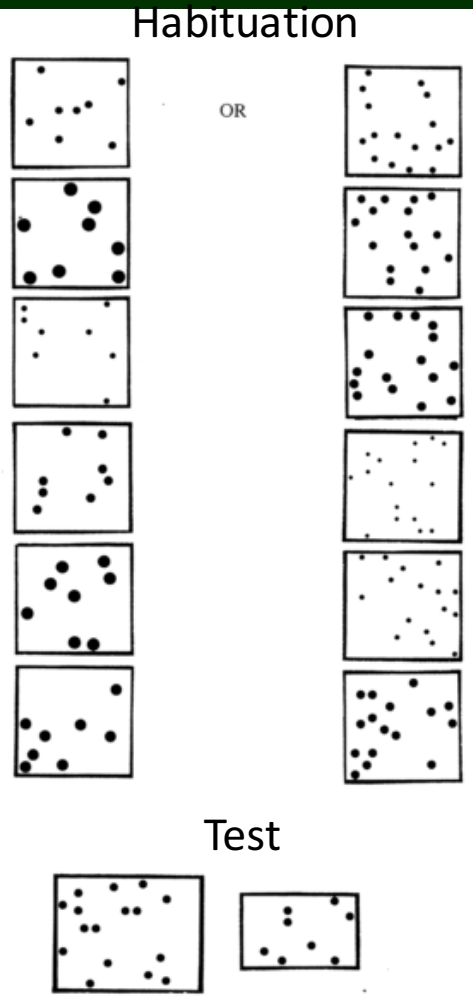
- Well-characterized, ratio-dependent
- Parallel detection, local adaptation (Burr & Ross 2008)
  - Intermediate representation with spatial receptive fields
- Babies can distinguish large approximate numbers

Birth: **1:3** (Izard et al., 2009)

6 months: **1:2** (Xu and Spelke, 2000)

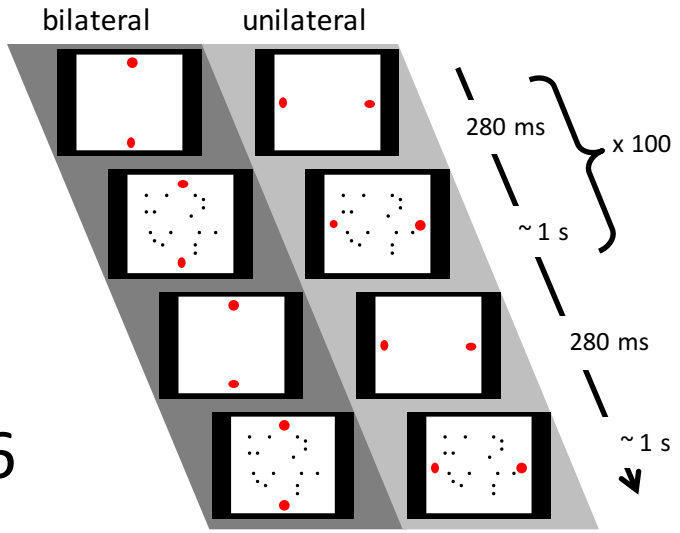
10 months: **2:3** (Xu & Arriaga, 2007)

Adult: **7:8** (Huntley-Fenner 2001, Mazzocco et al, 2011)



Do infants combine number across the midline?

1. Familiarization with 16 dots

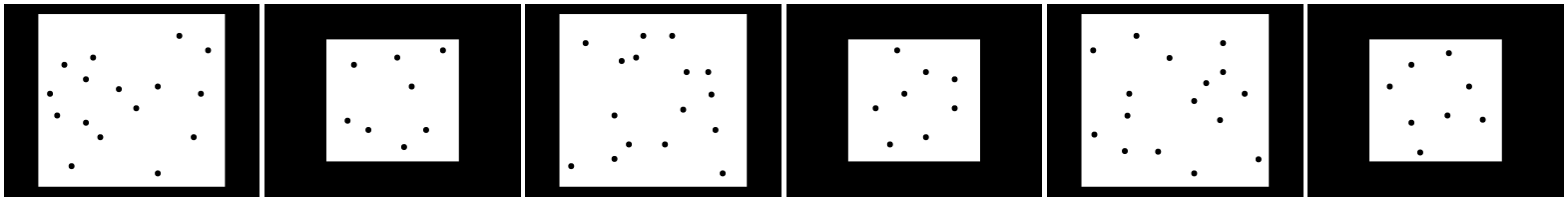


"Split-brain" prediction:

Unilateral: see 16 dots

Bilateral: see 8 dots

1. Test: Looking time to 8 and 16 dots. Measure until first continuous 1 second lookaway.

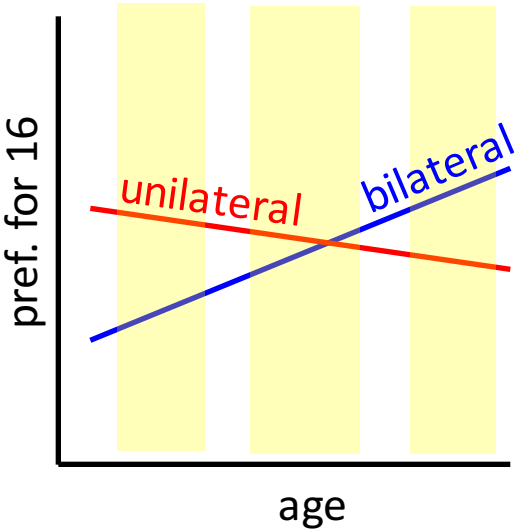


Looking time effect is hard to predict

Quality of the representation affects preference :

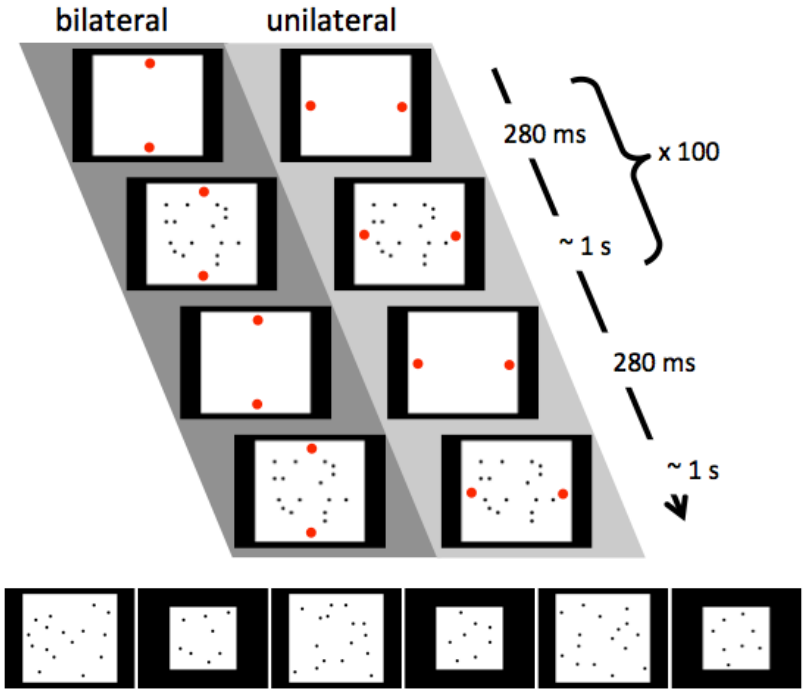
Familiarity preference	Novelty preference
Complex stimuli	Simpler stimuli
Fewer presentations	More presentations
Younger infants	Older infants

(Rose, Gottfried, Melloy-Carminar, & Bridger, 1982; Hunter, Ames, & Koopman, 1983; Richards, 1997; Roder, Bushneil, & Sasseville, 2000; Houston-Price & Nakai, 2004; Kidd, Piantadosi, & Aslin, 2012)



We check for **any** effect of condition or condition * age on looking preferences.

Results: Babies say 16 + 0 ≠ 8 + 8



741 trials from 147 children, ages 10 – 19 months
additional 92 children excluded/fussed out

Exclude individual trials:
- No valid lookaway
- Looking time < 1 s
- Child is fussy during lookaway
- Parent talks to child during measurement

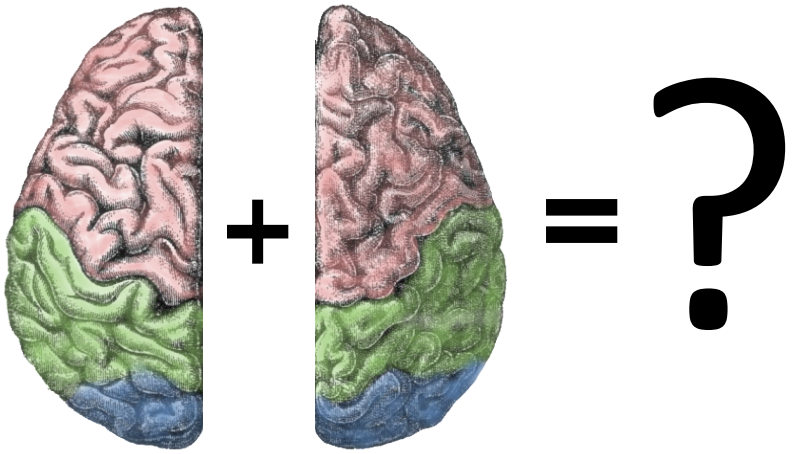
Hierarchical linear model on individual trial looking times:

LogLookingTime ~ Age * Bilateral * Stimulus16 + (TrialOrder – 1|Child) + (1|Child)

predictor	B	SE_B	t(731)	p
(Intercept)	0.64	0.03	24.16	0.000
Age (months)	0.02	0.01	1.16	0.248
Stimulus16	0.03	0.03	0.98	0.327
Bilateral	-0.01	0.04	-0.27	0.791
Age : Stimulus16	0.00	0.01	0.05	0.960
Age : Bilateral	0.00	0.02	-0.02	0.981
Stimulus16 : Bilateral	**0.12**	**0.04**	**2.76**	**0.006**
Age : Stimulus16 : Bilateral	**0.00**	**0.02**	**-0.07**	**0.944**

F(2, 731) = 3.82, p = 0.022

Next questions (in an ideal world)



- Is this really a novelty preference?
  - Compare 16 + 0 vs. 8 + 0
  - Do infants also show a preference for novel numerosities in the opposite direction (4 and 16 after seeing 8, 8 and 32 after seeing 16?)
  - Does more familiarization lead to stronger preference?
- By what age (and how fast) does integration occur?
  - What's represented in the process?
- Is the timing of integration linked across domains?
  - Comparing shapes, colors across hemifields
  - Use of language, other lateralized abilities

Collecting *these* data took 3 full years.

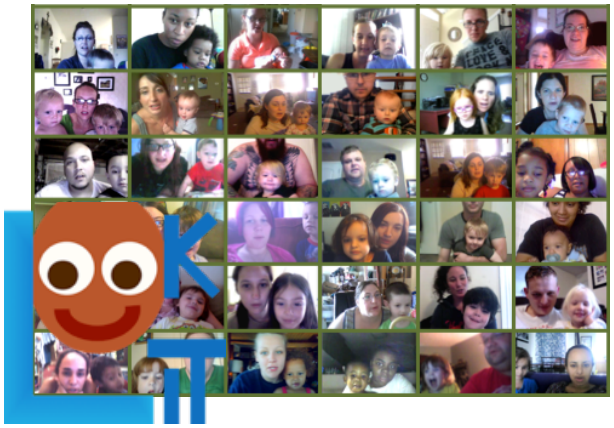
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Why put studies online?

Reducing practical constraints on the questions we can ask…

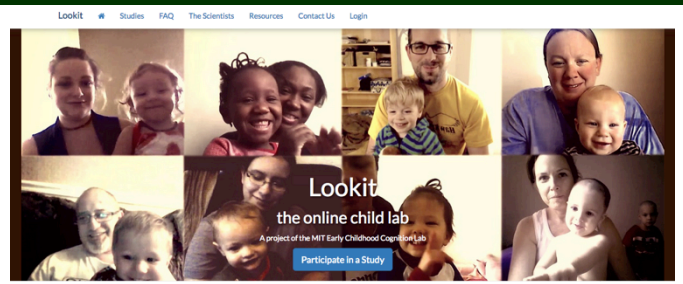
- Larger sample sizes
- Longitudinal designs
- Special populations
- Behavior at home
- More representative samples

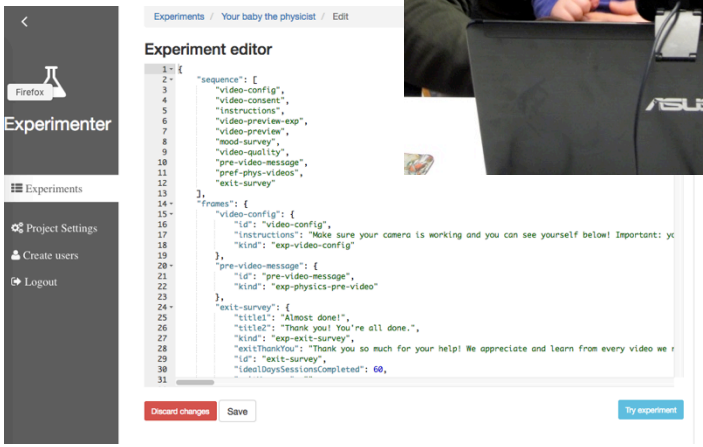
Designing for:

- Accountability and replicability
- Science outreach (“citizen science”)
- Bringing together computational, clinical, cognitive, educational approaches

How it works: https://lookit.mit.edu

- Parents register, complete child profiles, find and do studies in the web browser
  - Webcam video recording
  - No scheduling
  - No software download
- Researchers define & control studies from an 'Experimenter' platform
- Prototype site: ad-hoc study code, no experimenter interface
- Recruitment: early test studies used Mechanical Turk, currently volunteers





"I have read and understand the above information. I am this child's legal guardian and we both agree to participate in this study."

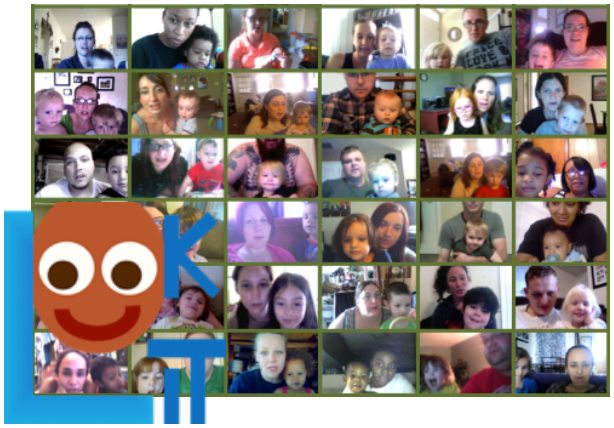
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Test studies on Lookit prototype

Intuitive probability

Age range: 11-18 months.

Infants have rich expectations about how the world works. Can they estimate how probable a completely new event is? In this study your child will see some 'lottery drawings' of shapes, for instance a yellow ball coming out of a mostly-blue container. We're interested in which outcomes he or she finds most surprising.

Learning new verbs

Age range: 2-year-olds (24-36 months).

In this experiment, your child will watch several clips of conversations that introduce a new verb, followed by side-by-side movies of experimenters pantomiming two different verbs. We are interested in how children use information from the conversations to figure out which action matches the new verb.

Learning from others

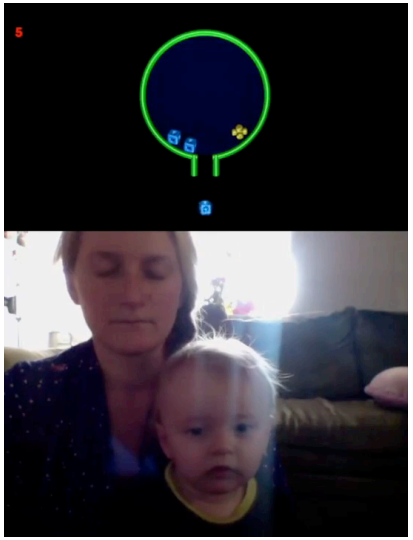
Age range: 3-, 4-, and 5-year-olds.

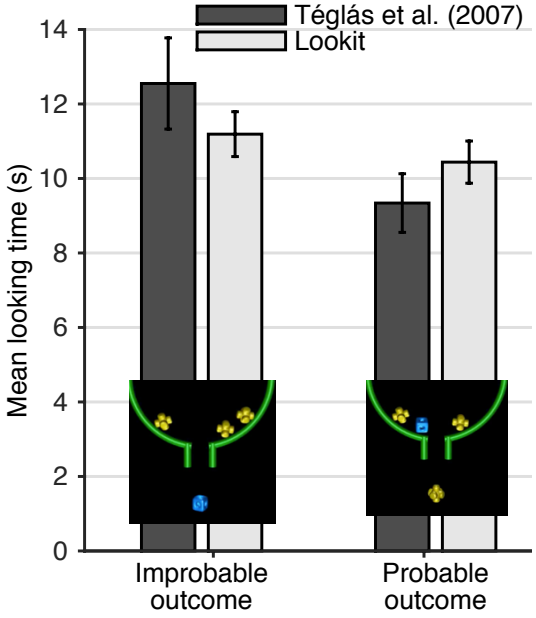
How do children filter through many sources of information to learn robustly from other people? In this study two women will name familiar objects with varying accuracy. We're interested in how your child chooses which one to trust when the same women name new objects as well.

Scott, K., & Schulz, L. (2017). Lookit (Part 1): A new online platform for developmental research. *Open Mind*.
Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*.

Study 1: One-shot probability

- Replicating Teglas et al. (2007): 12-month-olds look longer when the minority object exits a lottery container
- Four warmup (2:2) and four test (3:1) trials
- Parents close their eyes for the outcomes
- Measure **looking time** until first continuous 1 s lookaway on each trial.
- **49 children included, 11-18 months**.



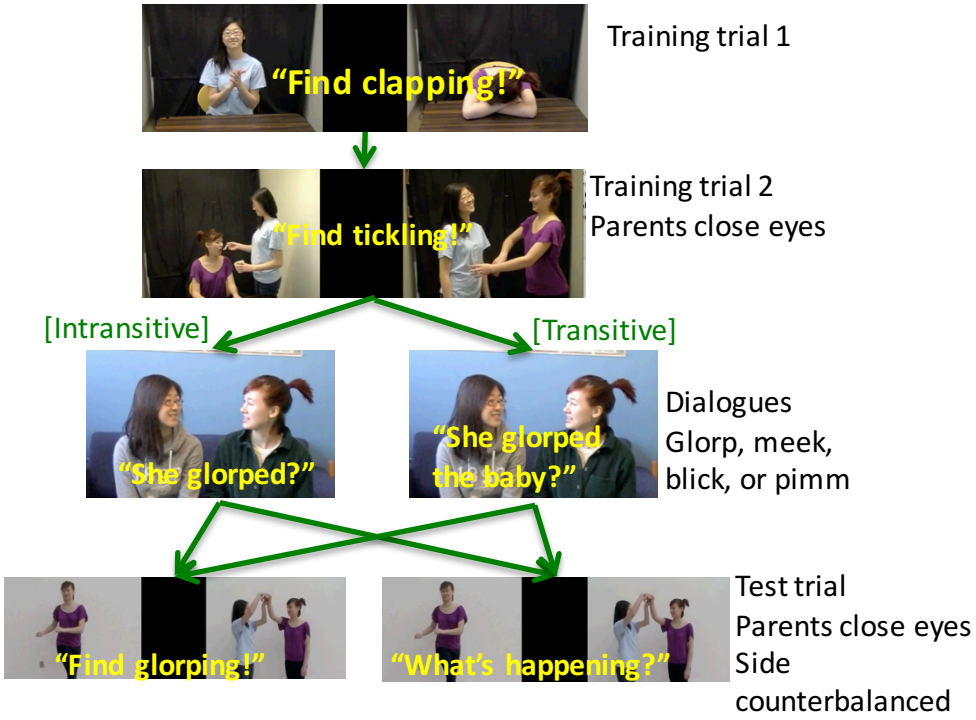


	B	SE B	p	95% CI
Intercept	13.08	0.94		[11.23, 14.94]
Improbable outcome	**0.62**	**0.66**	**0.35**	**[-0.68, 1.93]**
Trial order (1 - 4)	-0.86	0.29	0.003	[-1.44, -0.29]

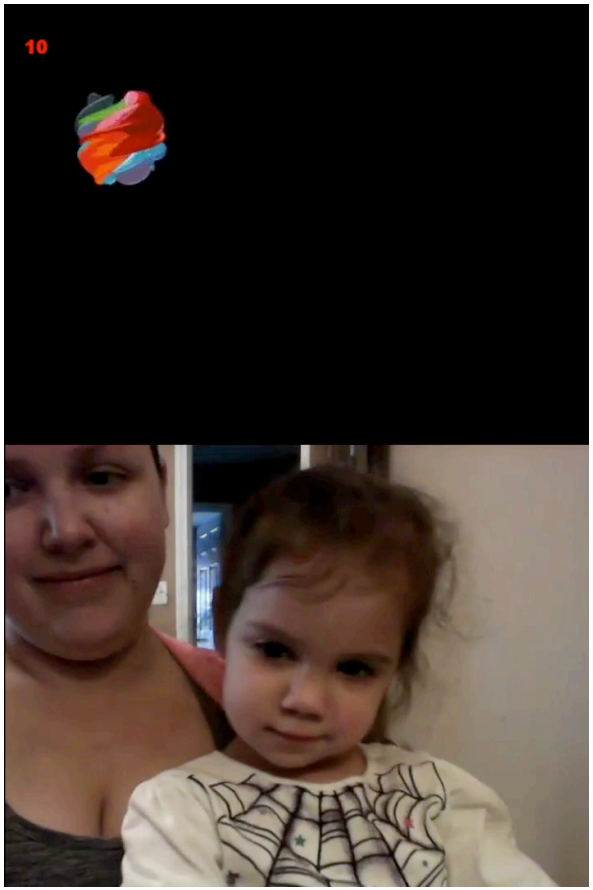
Difference in looking time, Téglás et al.: **3.21 s [0.36, 6.1]** (replicated internally)

Study 2: Syntactic bootstrapping

- Replication of Yuan & Fisher (2009): 2-year-olds use combinatorial information about a new verb to interpret its meaning
- Those who heard a transitive verb introduced looked longer than those who heard an intransitive verb at two-participant videos when asked to "find blicking!"
- No effect of transitivity when asked "what's happening?" instead

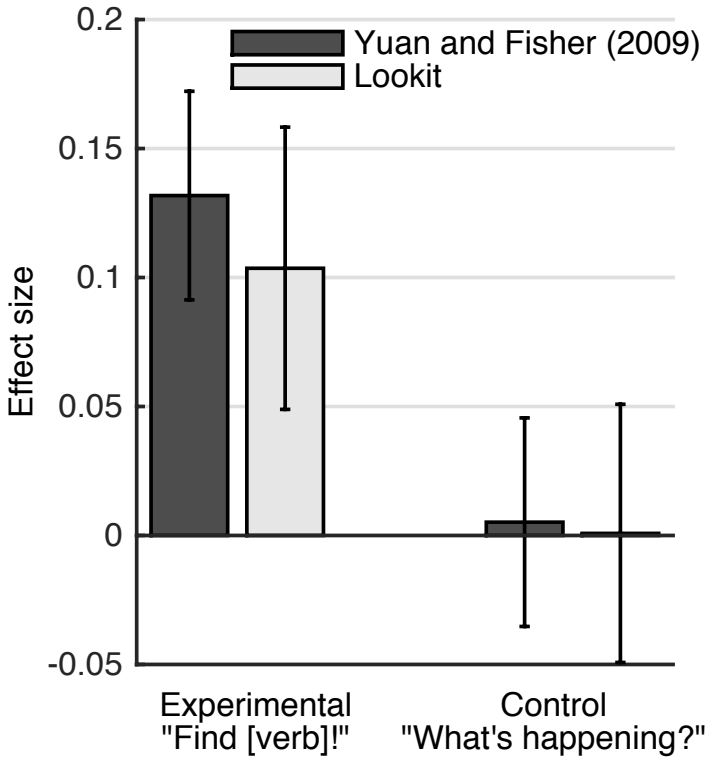


Study 2: Syntactic bootstrapping



Preferential looking, 67 toddlers (24-36 months)

How much does verb type affect action preference?



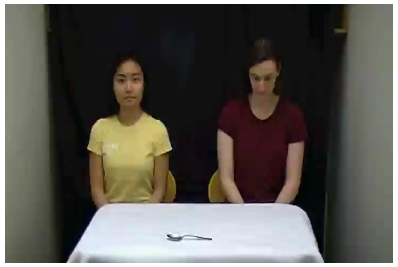
Fraction looking time to two-participant video
~ experimentalCondition * transitiveVerb + stimulusSet
Weighted by variance estimated from practice trials

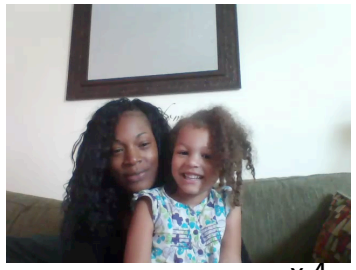
Study 3: Trust in testimony

Replicating Pasquini et al. (2007)'s "trust in testimony" work: Preschoolers use past reliability of informants to to identify who was more accurate and endorse her labels for novel objects.

- **Verbal responses**
- **148 preschoolers (3-4 years old) included**
- Informant accuracy conditions:
  - 100% vs. 0%
  - 100% vs. 25%
  - 75% vs. 0%
  - 75% vs. 25%

1. Familiar objects

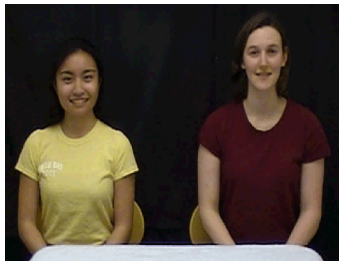




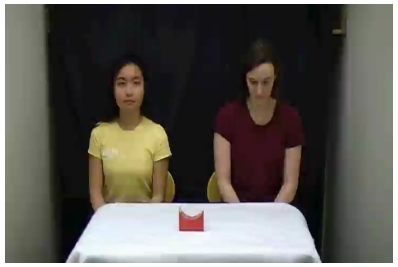
x 4

Counterbalance who's more accurate and when incorrect answers happen

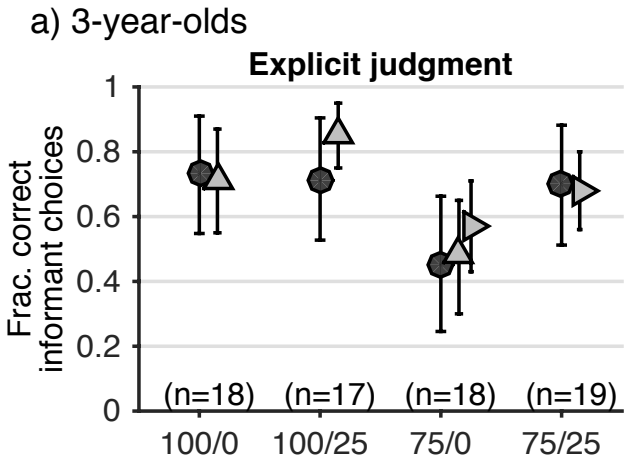
2. Explicit judgment



3. Novel objects



Study 3: Trust in testimony

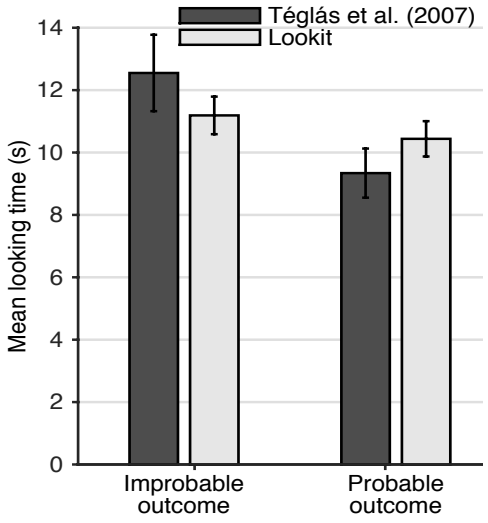


Lookit
Pasquini et al. (2007), Exp 1
Pasquini et al. (2007), Exp 2

Test study results

Oneshot probability: 12-month olds look longer at less probable events, even without prior experience

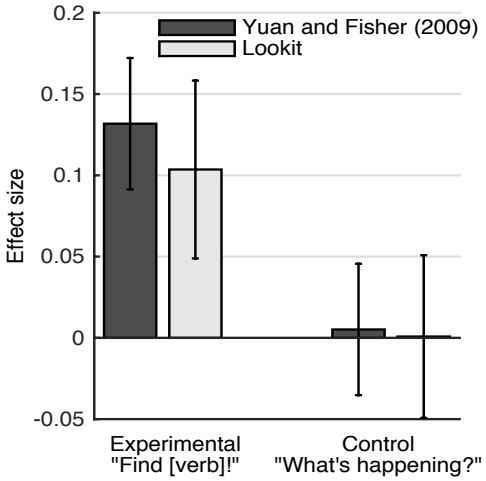
Looking time, 49 infants



Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. Proceedings of the National Academy of Sciences, 104(48), 19156-19159.

Syntactic bootstrapping: 2-year-olds use combinatorial information about a new verb to interpret its meaning (semantic bootstrapping)

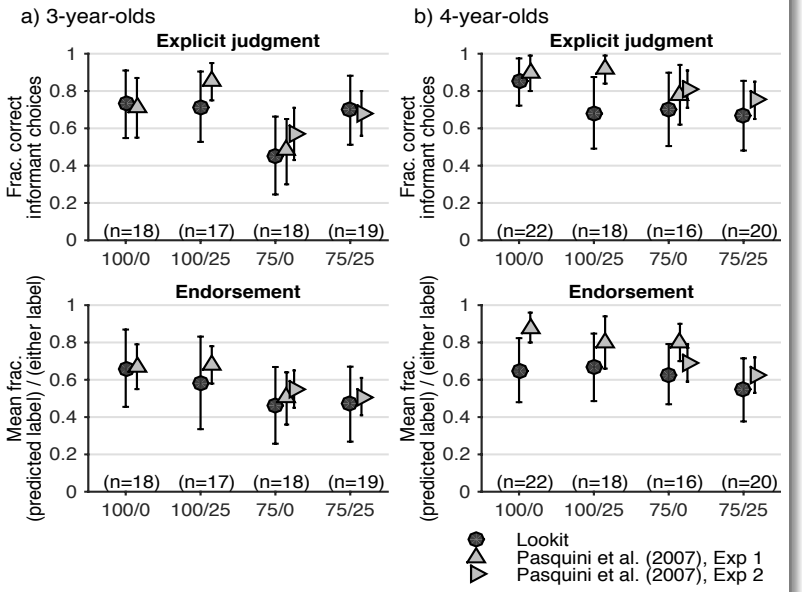
Preferential looking, 67 toddlers



Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619-626.

Trust in testimony: Preschoolers use past reliability of informants to choose whose labels of novel objects to identify who was more accurate and endorse her labels

Verbal responses, 148 preschoolers



Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. Developmental psychology, 43(5), 1216.

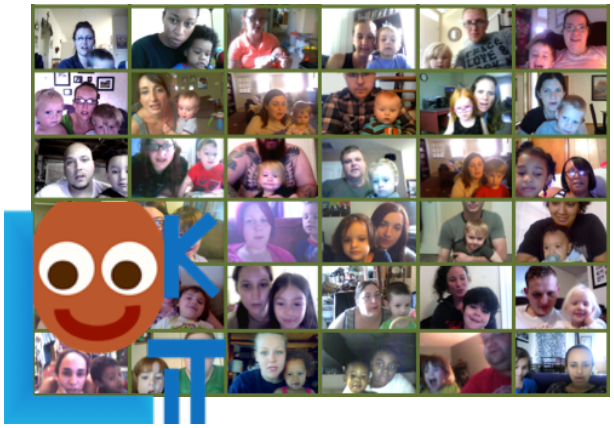
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield

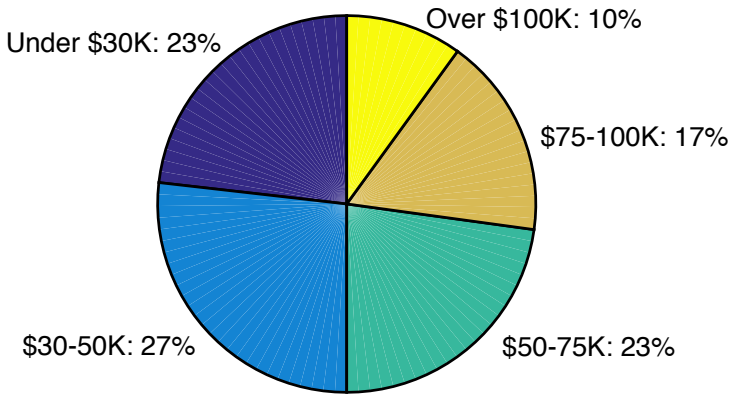


Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Online participants are more diverse

Family income

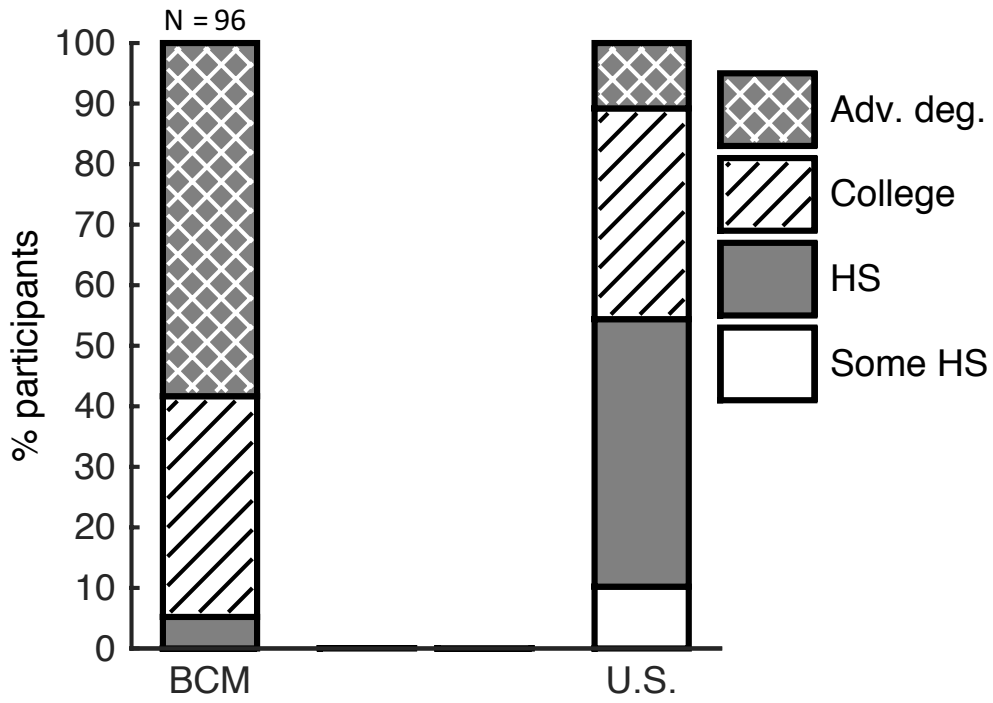


Race of child & Hispanic origin

	Lookit	US population
Hispanic	9.6	17.4
White	79.6	77.4
Black	7.1	13.2
Asian	1.6	5.4
American Indian, Alaska Native, Native Hawaiian, or Pacific Islander	0.4	1.4
2 or more races	11.3	2.5

- Based on 572 surveys (75% response rate) from families who gave valid consent to participate in at least one test study
- Almost all recruited from MTurk
- Languages:
  - 8% at least bilingual
  - 21 Languages represented: Arabic, ASL, Azerbaijani, Bengali, Chinese, Czech, French, German, Hindi, Hungarian, Japanese, Korean, Portuguese, Romanian, Russian, Spanish, Tagalog, Telugu, Turkish, Urdu, Vietnamese
- Still mostly moms (84%)

...and less over-educated



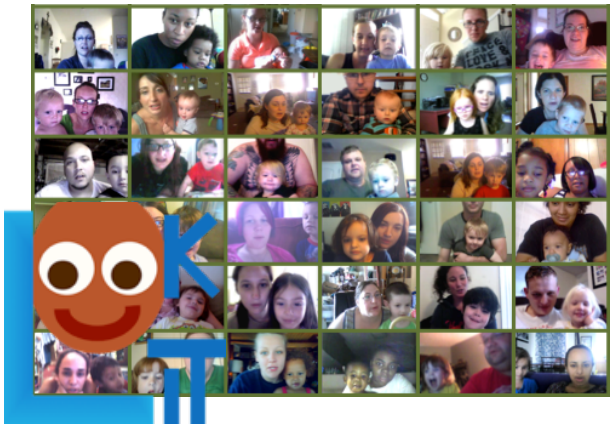
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Coding looking time works!

Looking time (looking vs. not) blind coders agree 95% of the time (N = 63 children, SD = 6%), mean absolute difference in LT is 0.8 s (SD = 1 s)

Coding preferential looking works!

Looking time (looking vs. not)

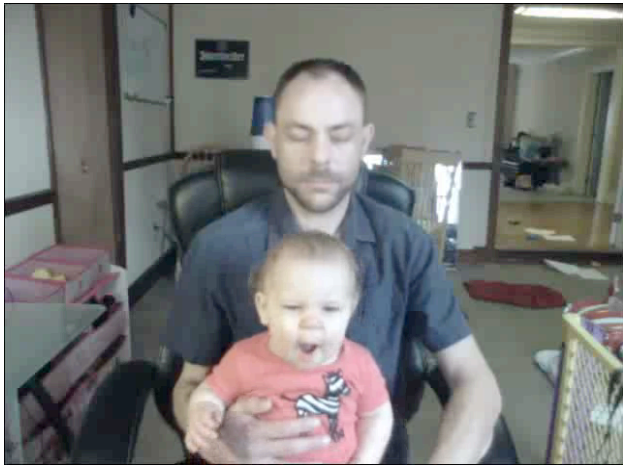
blind coders agree 95% of the time (N = 63 children, SD = 6%), mean absolute difference in LT is 0.8 s (SD = 1 s)

Preferential looking (right vs. left)

Blind coding done close to real-time; mean disparity between coder's judgments of looking time to left/right is 4% of trial length (SD = 2%, N = 140 children)

But what about data quality?

One of our postgraduate students kindly agreed to provide reliability coding for another postgraduate student, who had found null results in a visual habituation study. These two students had both been trained on lab procedures, including criteria for eliminating untestable infants. Despite this, their working definitions of excessive "fussiness" were apparently different, because as the reliability coder watched the videotapes she was surprised to see that a number of the infants completed their looking task while simultaneously squirming in their seats, making whining noises and in some cases overtly weeping.



For all looking studies, two coders recorded for each clip:

fussiness
distraction
parental interference,

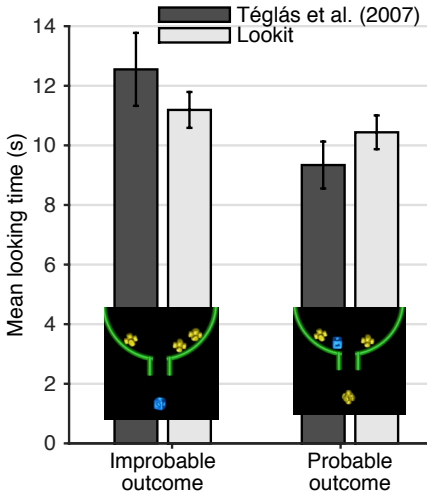
Slaughter, V., & Suddendorf, T. (2007). Participant loss due to "fussiness" in infant visual paradigms: A review of the last 20 years. Infant Behavior and Development, 30(3), 505-514.

Coders agree on "fuzzy" events

	Intuitive probability (113 infants)			Learning new verbs (141 toddlers)		
	Agreement	Cohen's κ	Frequency	Agreement	Cohen's κ	Frequency
Fussy (crying or trying to get out of parent's lap)	85%	0.55	13%	99%	0.42	1%
Actively distracted (lookaway caused by external event)	85%	0.40	10%	97%	0.37	2%
Parent's eyes open	98%	0.67	4%	97%	0.70	5%
Parent's eyes not visible	97%	0.35	4%	99%	0.91	7%
Parent peeks	90%	0.36	9%	90%	0.36	11%
Parent's eyes open briefly at very start of clip	77%	0.49	38%	93%	0.68	13%
Parent talks to child	98%	0.16	2%	95%	0.70	9%

Child attentiveness at home

Common concern: home is “noisier” than the lab…



Indeed, relatively frequent distractions for LT measurements (10%)

BUT…

- Looking times similar to those found in lab (~10 s)
- Excellent attention (>80% looking) to dialogue videos during preferential looking study

Parents cooperate online

- Most parents report that they did keep their eyes closed (63%, 67%) or that they tried (28%, 25%).
  - "I always kept my eyes closed except for once when my child was moving the computer with his feet."
  - "yes, and she reprogrammed my laptop smashing keys."
  - "Looked when he said uh oh, just to make sure he wasn't talking about his siblings"
  - "Well, I did until my 6 year old came in unexpectedly. Sorry"
- But never peeking is hard (just like in the lab): 22% of parents peeked during at least one test trial

Adapting looking measures

Advice – parent blinding:

- Have parents hold children looking over their shoulder, give practice trials if possible
- Keep parents in the loop: play audio to confirm everything's working
- Let parents see all the stimuli ahead of time

Advice – engagement:

- Use parents rather than recorded audio if possible
- Use methods robust to distractions
  - preferential looking over looking time
  - hierarchical linear models or Bayesian analysis to deal with missing trial

Can reliably code looking time & preferential looking from webcam!

I admit I got suspicious when I had to close my eyes because [Child] would start responding out loud to whatever was on the screen. But I kept them closed! Everything ran great.

I peeked a tiny bit but was careful to make no reaction. Even though I am sure that you are a reputable research group, as a parent, I feel that it is my responsibility to monitor what she is watching and seeing.

I opened them once to make sure my child was being shown appropriate things (people are strange out there!)

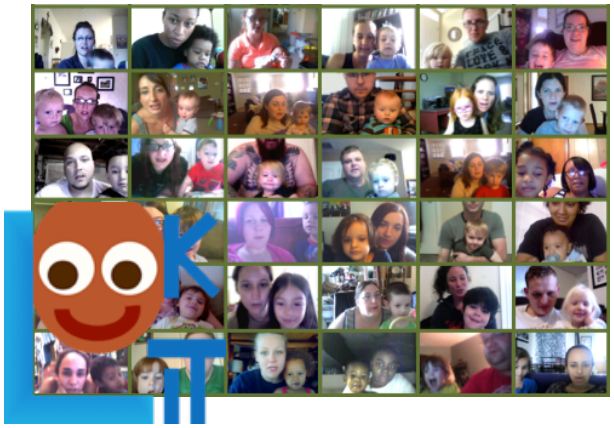
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

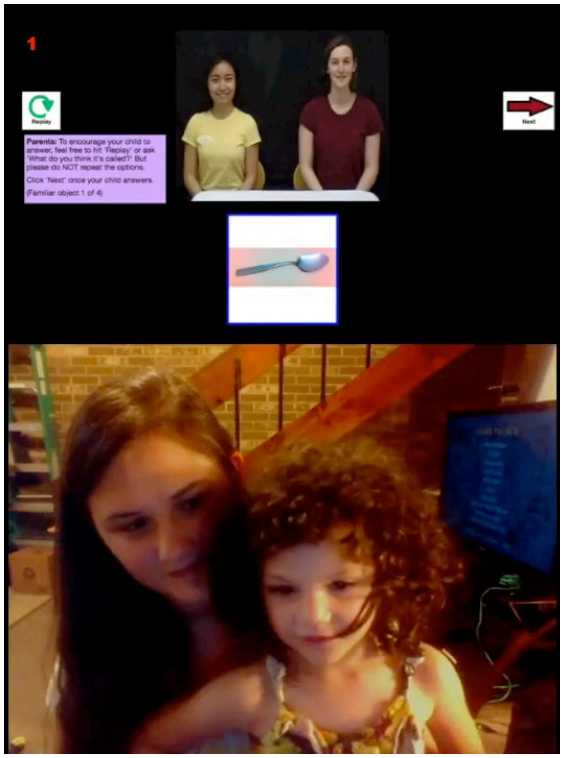
- How it works
- Example use cases
- A vision for Lookit

Collecting verbal responses

- Parent and child speech transcribed
- Parent interference and interaction types recorded
- Only use responses before any parent interference (8% of trials). Examples: child [parent]

  - What do you think, mom? [I want your answer, not what mommy thinks.] **Red**.

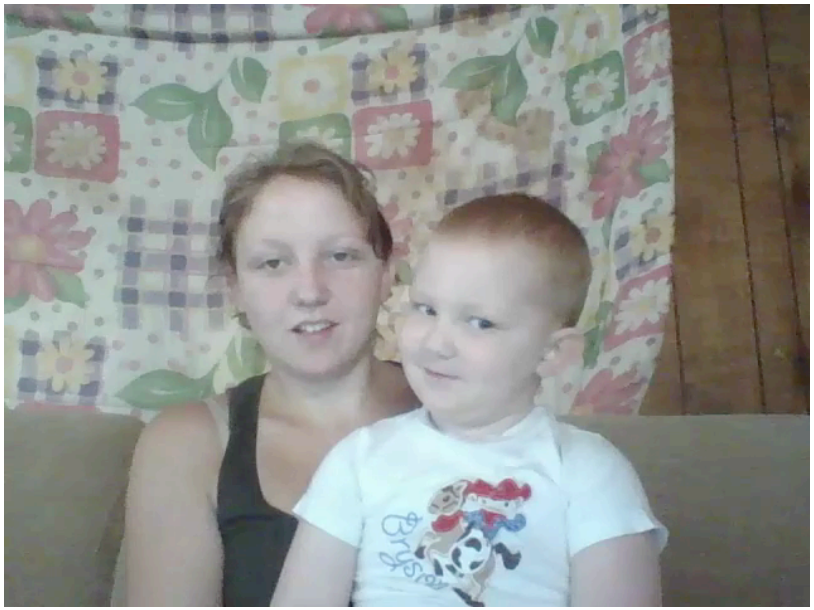
  - I don't know how to say those words. [Try it, which one are you trying to say? What do you think it is?] I don't know how to say those words. [Just try it, you're doing good. Try to pronounce it.] I can't, I can't though. [Okay. Do you want to listen to it again?] Yeah. [REPLAY.] A **toma**? [Then say it again...] A toma!



Verbal responses

Recorded questions asked to children, with parent instructions for prompting (234 preschoolers with usable video)

“More natural behavior” at home



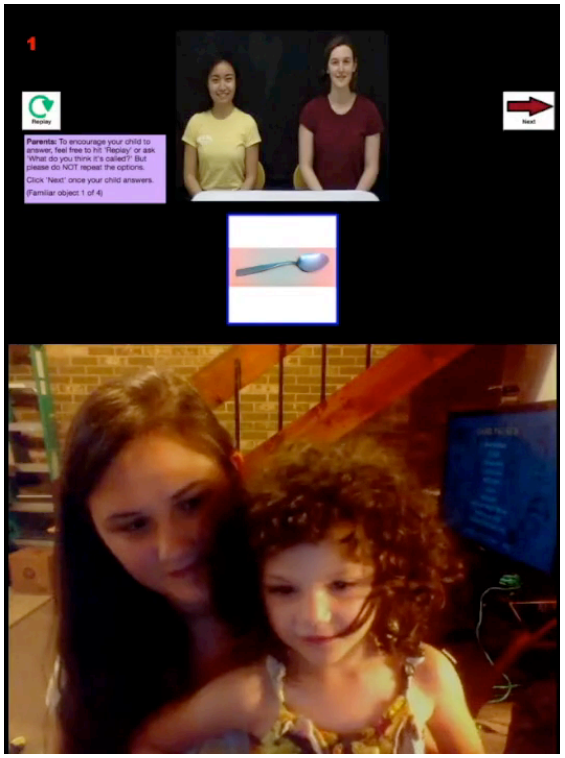
We rely on parents’ help!

Adapting verbal response measures

Advice

- Encourage responses
  - Involve the parent where possible
  - Give computer-based contingent responses
- Show parents exactly what to do/what to avoid
  - Include video examples
  - Give practice trials before data is critical
- Design studies to be robust to non-responses
  - Many short, independent questions
  - Analysis to use whatever you get

Can collect verbal responses online!



Verbal responses

Recorded questions asked to children, with parent instructions for prompting (234 preschoolers with usable video)

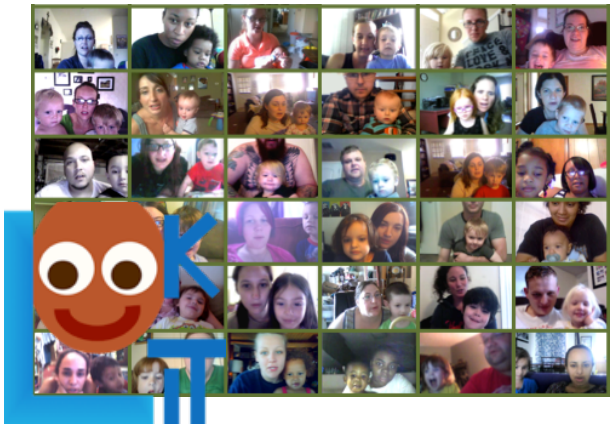
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Yield is lower online

Percentages	Looking time	Pref. looking	Verbal resp.	Total
Invalid consent	19	16	15	**16**
Out of age range	7	4	7	**6**
Unusable video	32	38	20	**29**
Video is okay…	42	43	59	**49**
Actually included	*18*	*20*	*37*	***26***
Total	**100**	**100**	**100**	**100**

50% no video
15% incomplete
20% low framerate or too dark/blurry
1% eyes generally not visible on screen

- **Mostly** lower due to online-specific issues early in (or before) data collection

Exclusion rates	Looking time	Pref. looking	Verbal resp.
Lookit	56% (63/112)	52% (73/140)	37% (86/234)
Original lab study	50% (20/40)	12% (13/92)	6% (6/98)

- Time cost is **still** lower to collect data online

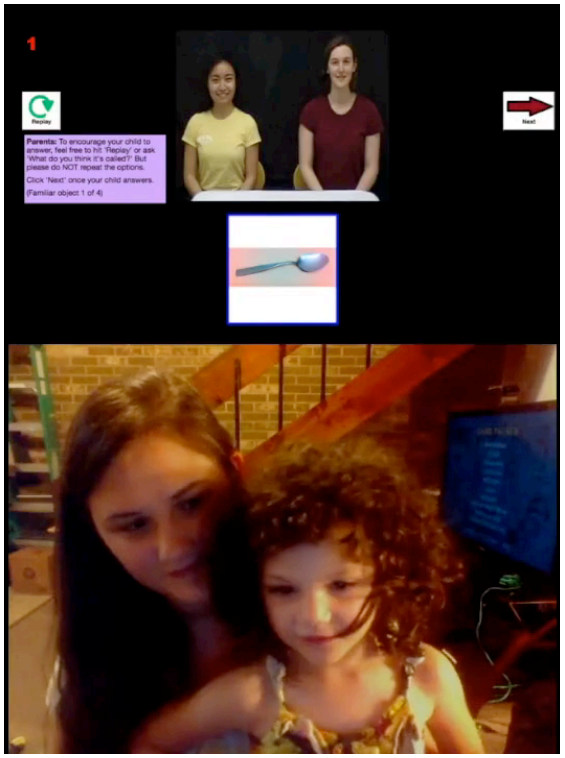
Data we can collect

Looking time (looking vs. not)

blind coders agree 95% of the time (N = 63 children, SD = 6%), mean absolute difference in LT is 0.8 s (SD = 1 s)

Preferential looking (right vs. left)

Blind coding done close to real-time; mean disparity between coder's judgments of looking time to left/right is 4% of trial length (SD = 2%, N = 140 children)



Verbal responses

Recorded questions asked to children, with parent instructions for prompting (234 preschoolers with usable video)

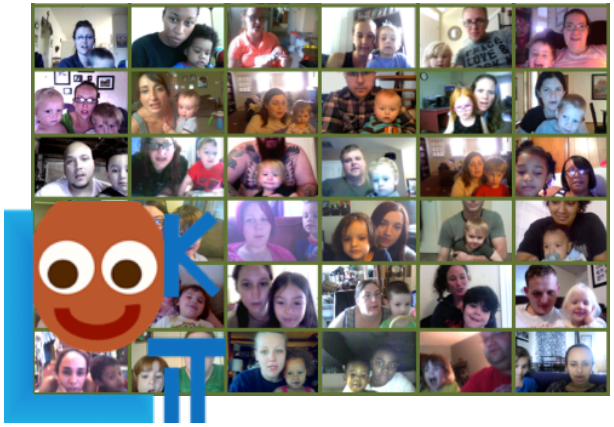
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield

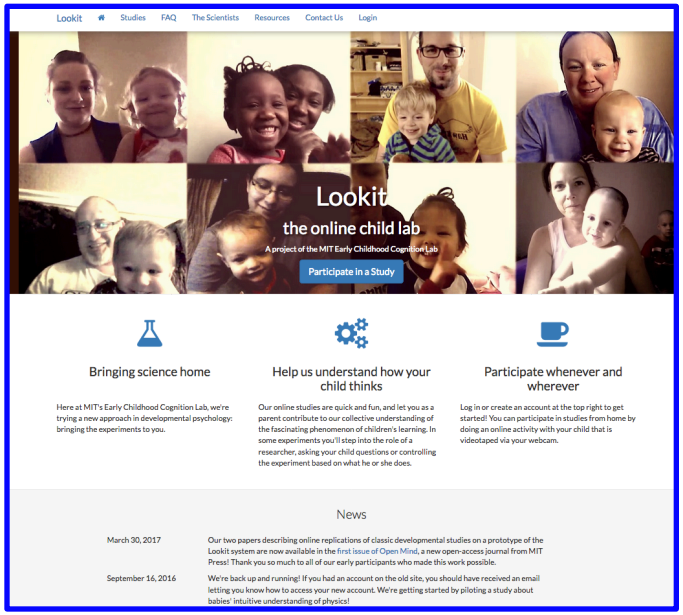


Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

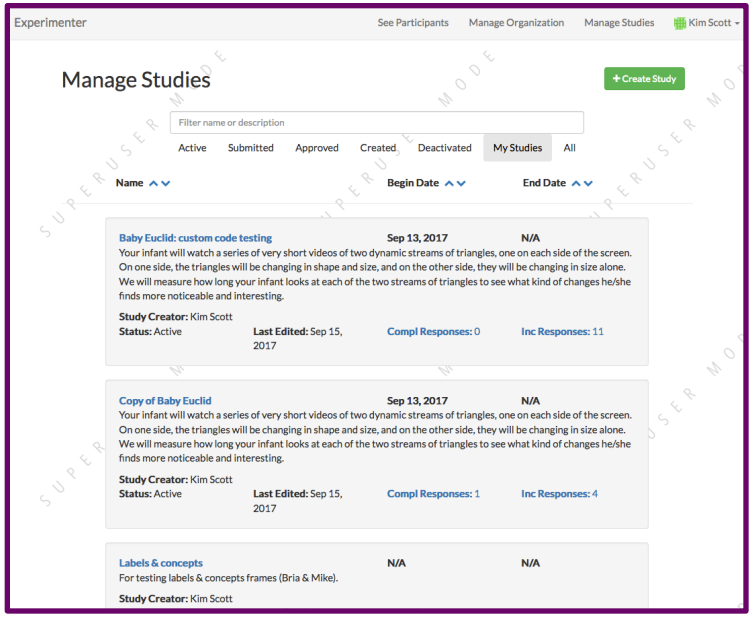
Towards a multi-lab platform

Lookit



Participate in studies

Experimenter



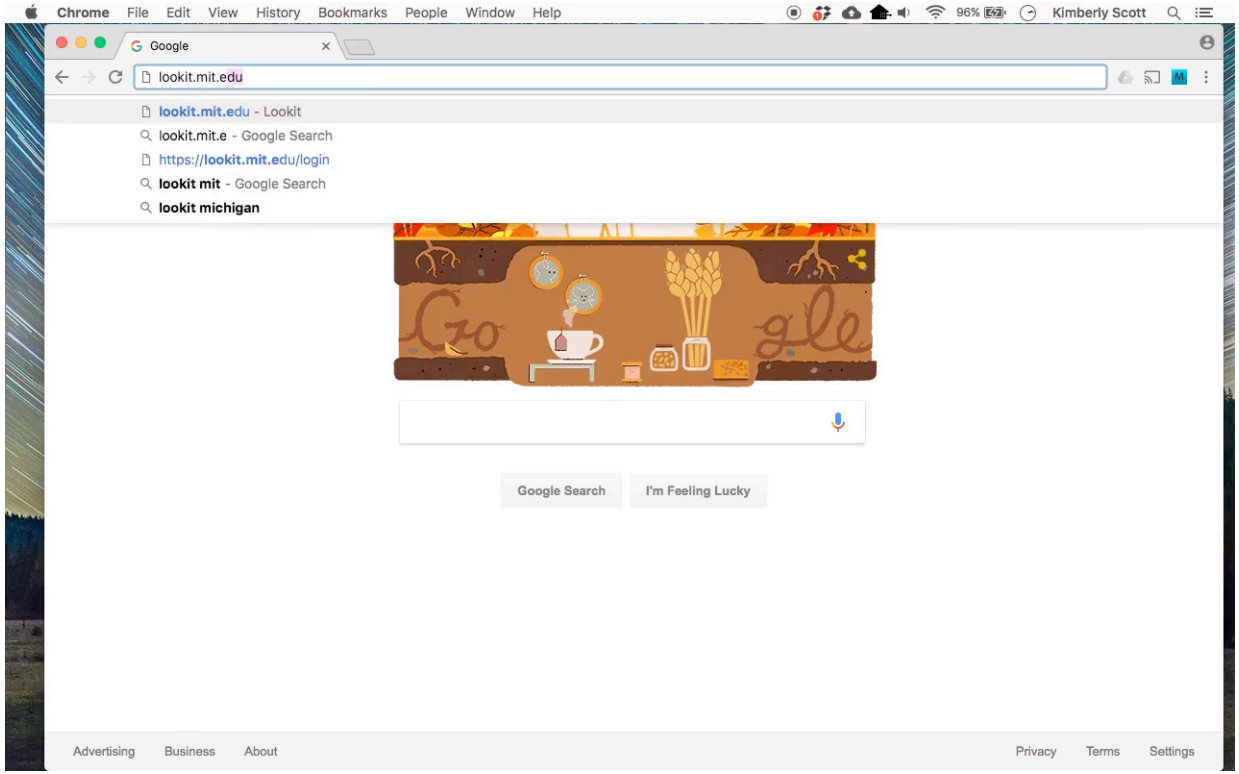
Design, manage, & collect data from studies

Code available at: https://github.com/CenterForOpenScience/
- ember-lookit-frameplayer
- lookit-api
- experimenter
- exp-addons

Participant experience

Account

- Any family can create an account
- Register any number of children
- General demographic survey
- Email preferences



Study participation

- Select from list of studies
- Webcam setup, verbal consent
- Choose privacy level at end
- Previous sessions & any feedback shown

Researcher experience

Permissions

- Researchers given access to Experimenter by admin
- Manage / view data from own studies only

Managing a study

- Preview on Experimenter
- Submit study for approval by an admin
  Pause/stop/restart at any time
- Studies can be active but "non-discoverable"

Getting data

- Download JSON/CSV data via Experimenter
- Download data about own participants only
- Download individual or bundled videos
- Programmatic access via API

Communication

- Email participants via Experimenter
- Actual email addresses hidden

Study protocols specified in JSON

JSON (**J**ava**S**cript **O**bject **N**otation) object example:

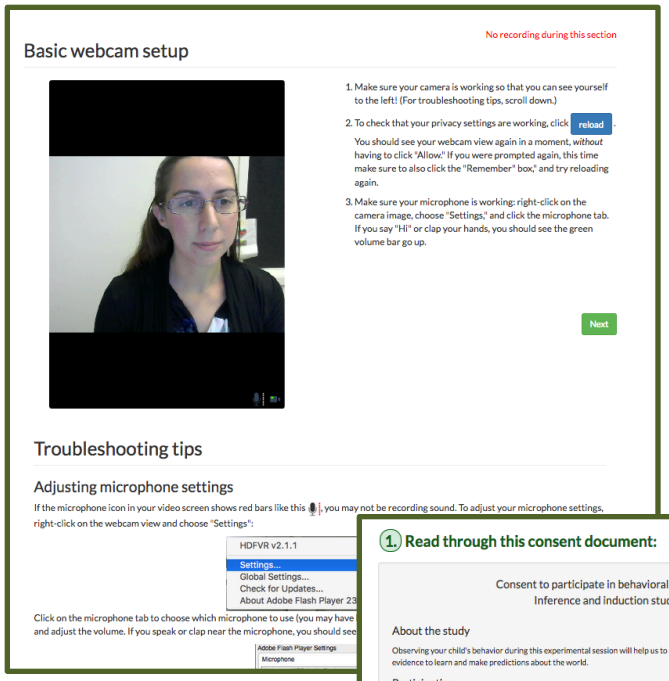
```
{
    “name”: “Jane”,
    “age”: 43,
    “favoritefoods”: [
        “eggplant”,
        “apple”,
        “lima beans”
    ],
    “allergies”: {
        “peanut”: “mild”,
        “shellfish”: “severe”
    }
}
```

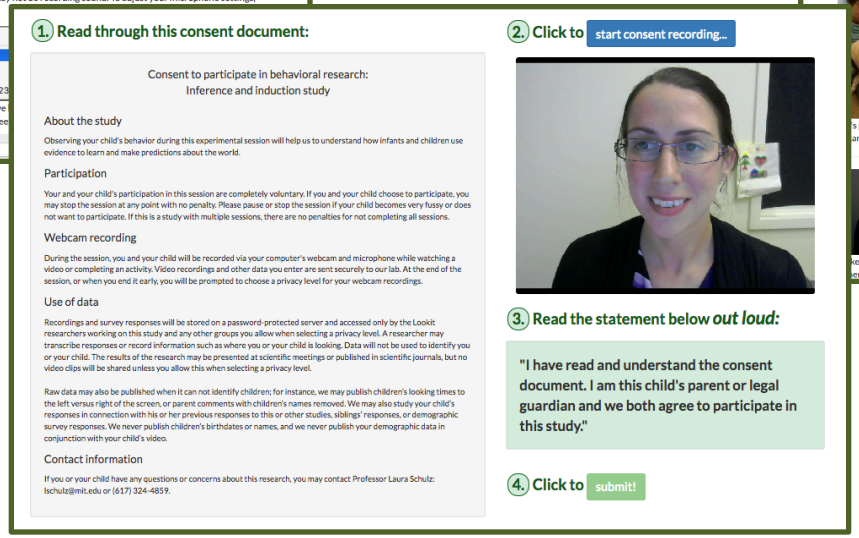
Specifying a study protocol

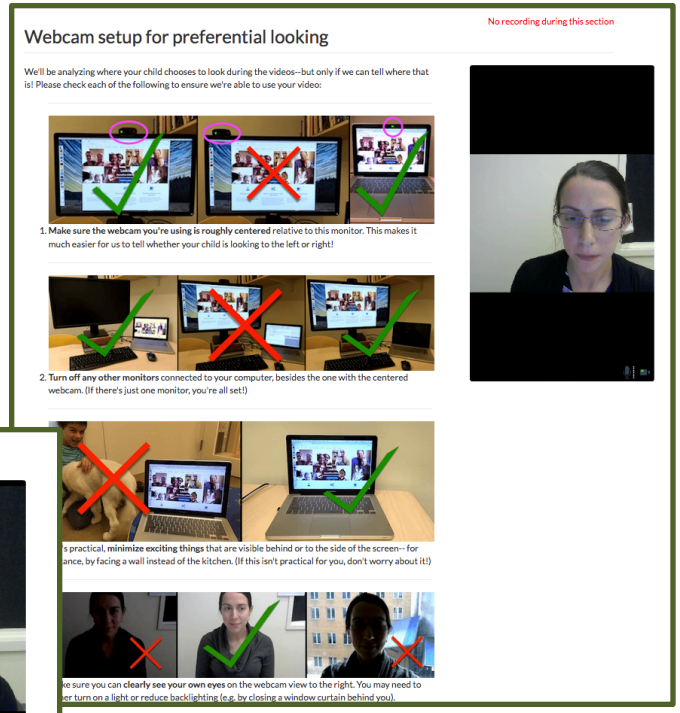
- Studies specified using JSON (JavaScript Object Notation), a simple text format
- Studies broken into "frames"
- JSON object needs a "frames" key and a "sequence" key

```
{
    "frames": {
        "my-consent-frame": {
            "kind": "exp-video-consent",
            "prompt": "I agree to participate",
            "blocks": [
                {
                    "title": "About the study",
                    "text": "Not a real study."
                }
            ]
        },
        "my-exit-survey": {
            "kind": "exp-lookit-exit-survey",
            "debriefing": {
                "title": "Thank you!",
                "text": "You participated."
            }
        },
        "sequence": [
            "my-consent-frame",
            "my-exit-survey"
        ]
}
```

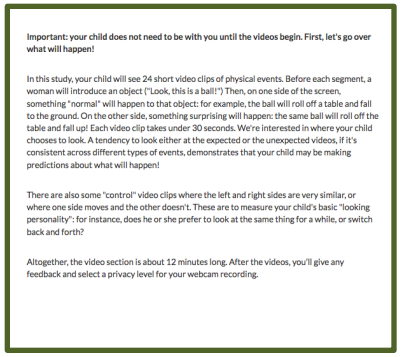
Common frames available

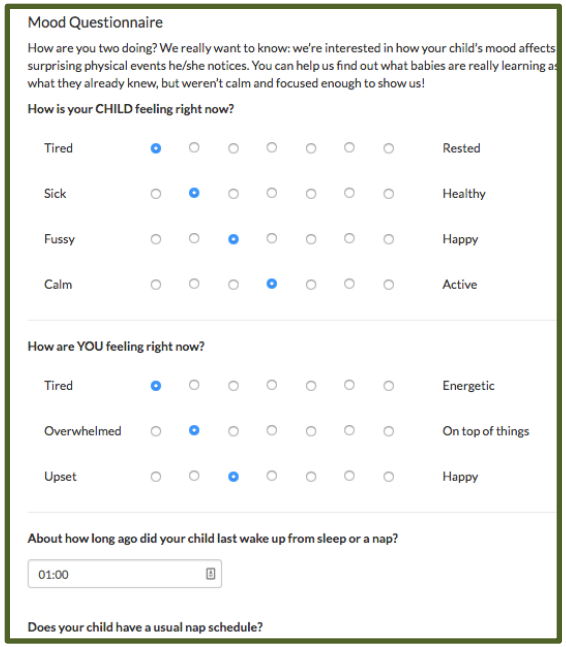


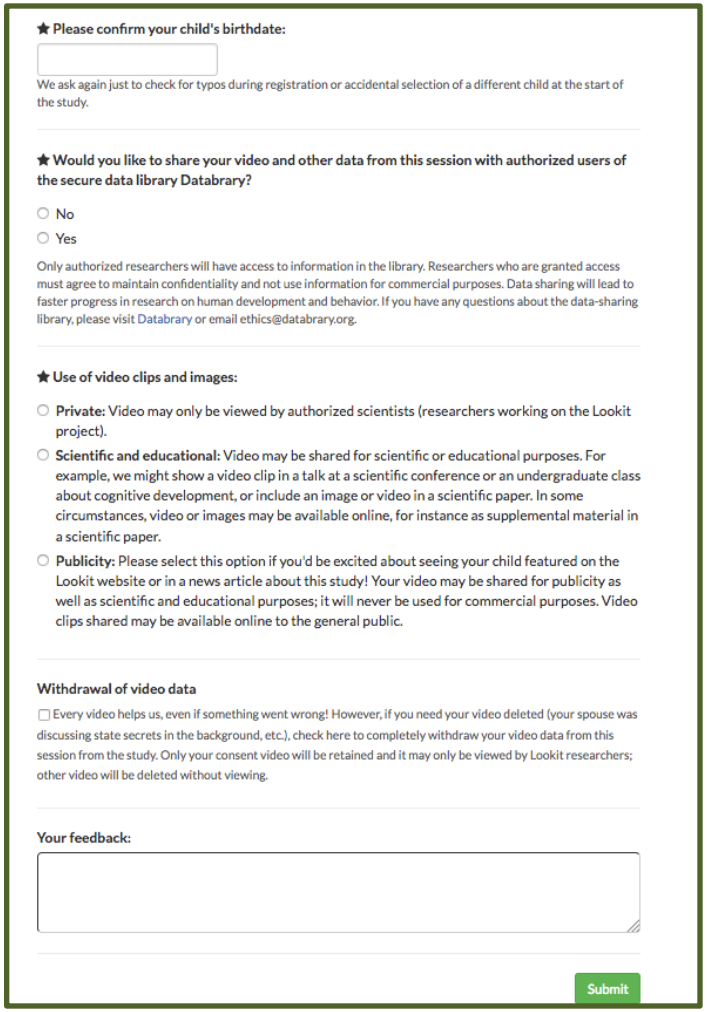




Common frames available



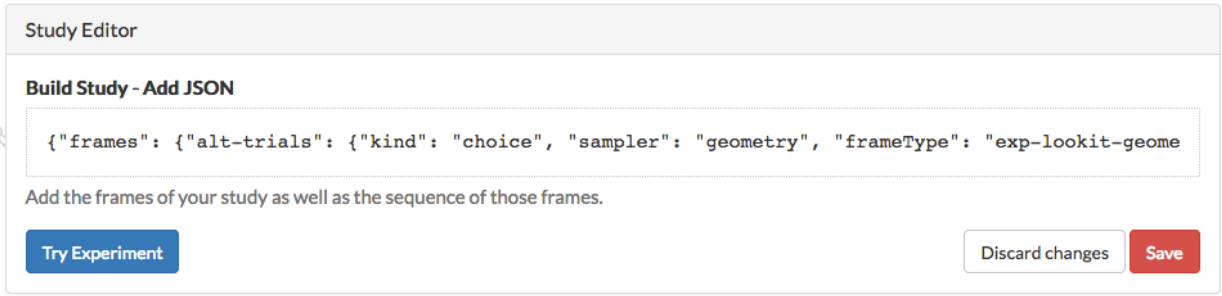


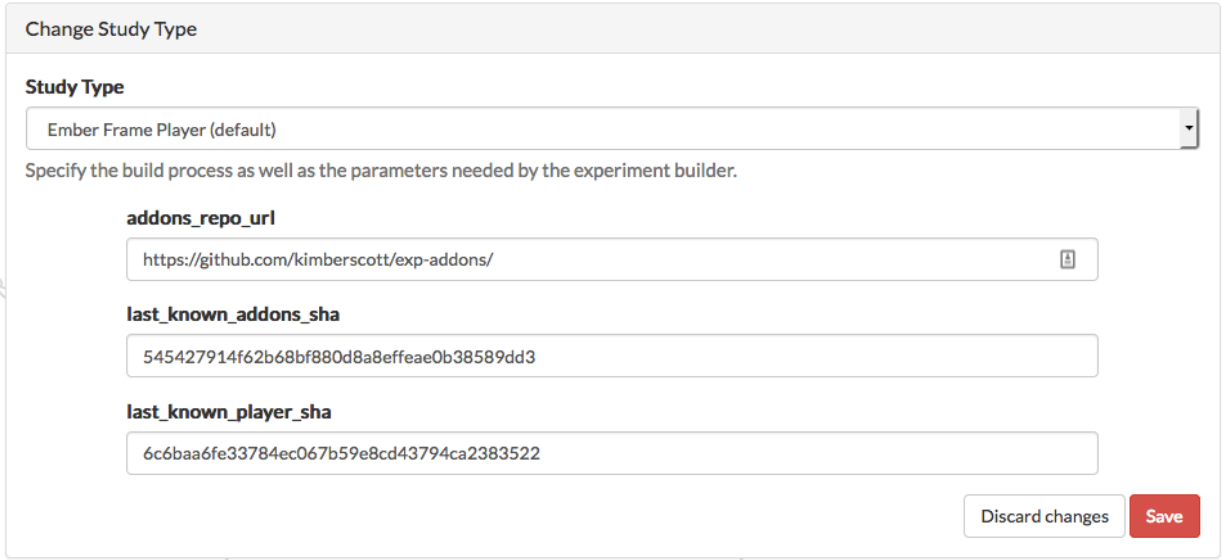


- "Randomizer" frames allow counterbalancing
  - Shuffle or select from frames
  - Arbitrary & nested designs
- Video recording behavior determined per frame
- Time of frame start + custom data saved to database throughout study

Or use your own custom code

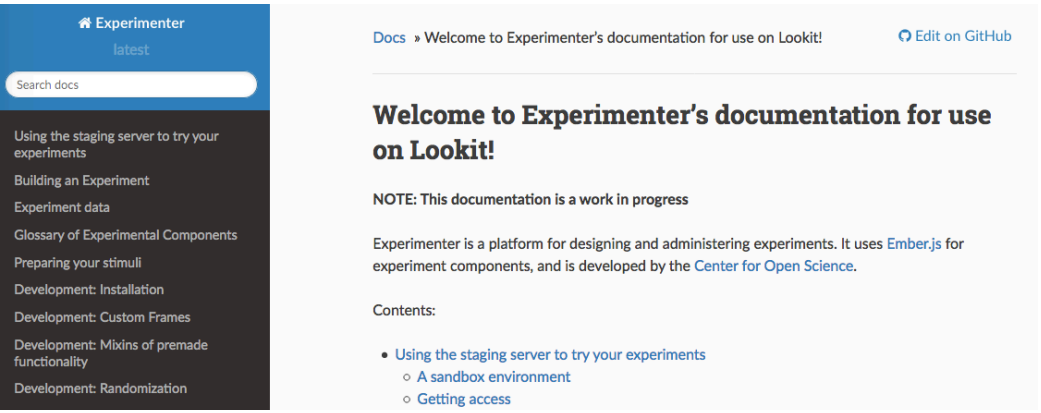
- Specify a commit of your own fork of exp-addons
- ALL code for your study is bundled up and deployed, separate from any other study code





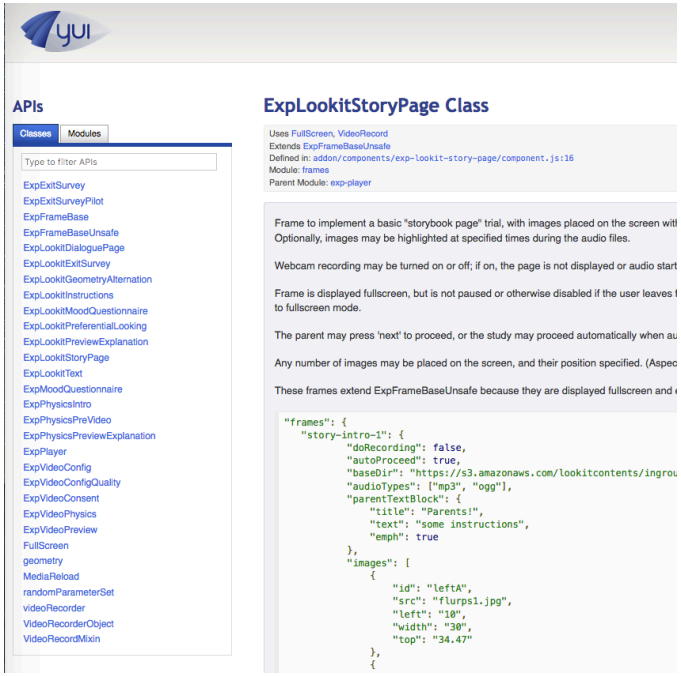
Documentation

Experimenter/study creation: lookit.readthedocs.io



Lookit API: lookit-api.readthedocs.io

Individual frames: centerforopenscience.github.io/exp-addons/ (automatically generated from comments)



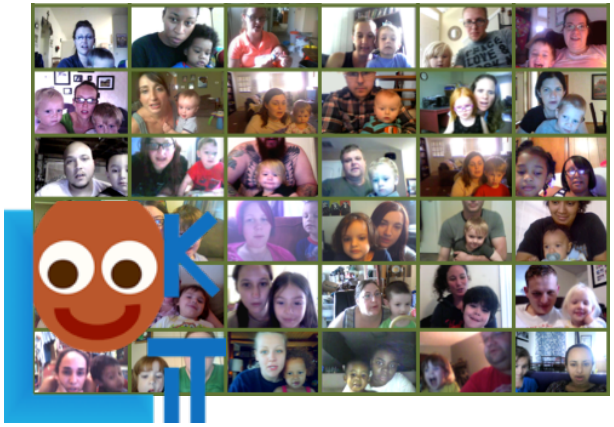
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Studies implemented on new platform

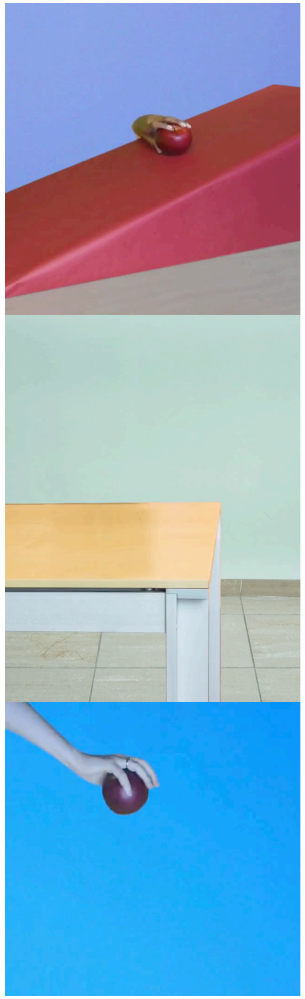
Round	Study	With…	Method
0 (our own study)	Dense measurements of intuitive physics judgments	[ECCL]	Preferential looking/VoE

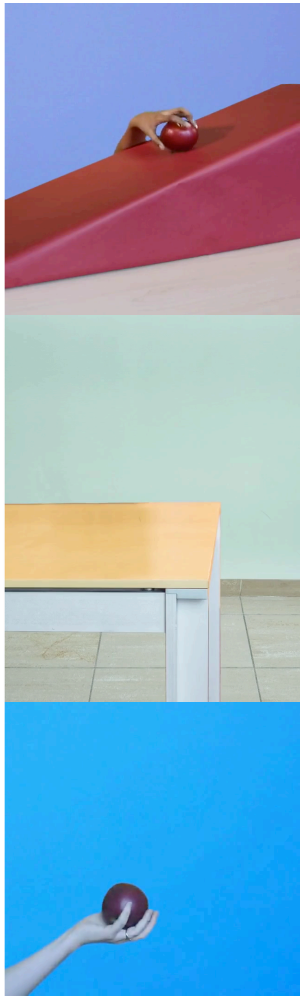
Dense measurements of intuitive physics judgments

Ongoing study: "Your baby, the physicist"

- 50 infants (4- to 12-months) participate online in a preferential looking study
- Each trial: introduce an object, then show a pair of more- and less-surprising events (loosely related to gravity, inertia, and support)
- Control measures about looking patterns interspersed
- Repeated measures: 24 trials per session, 15 sessions over 1-2 months.
- Also collect mood survey

Gravity comparisons



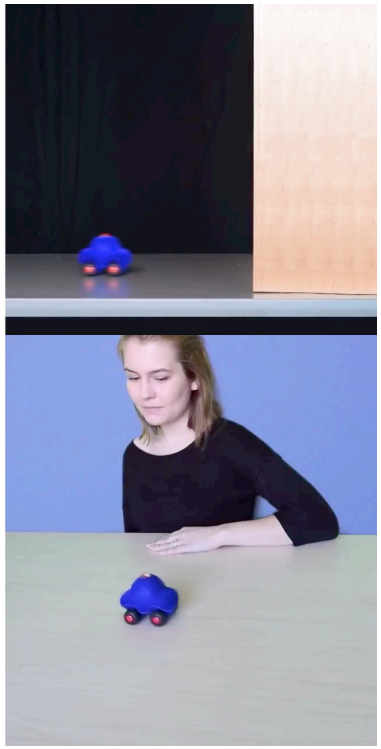


Dense measurements of intuitive physics judgments

Ongoing study: "Your baby, the physicist"

- 50 infants (4- to 12-months) participate online in a preferential looking study
- Each trial: introduce an object, then show a pair of more- and less-surprising events (loosely related to gravity, inertia, and support)
- Control measures about looking patterns interspersed
- Repeated measures: 24 trials per session, 15 sessions over 1-2 months.
- Also collect mood survey

Inertia comparisons



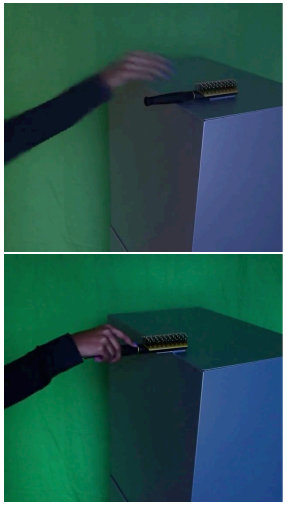
Dense measurements of intuitive physics judgments

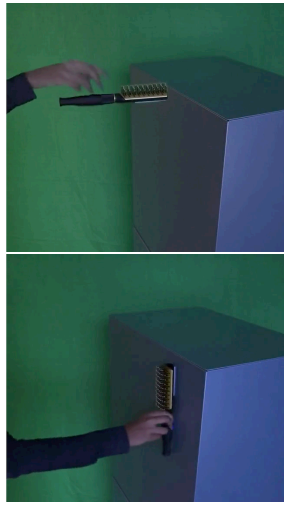
Ongoing study: "Your baby, the physicist"

- 50 infants (4- to 12-months) participate online in a preferential looking study
- Each trial: introduce an object, then show a pair of more- and less-surprising events (loosely related to gravity, inertia, and support)
- Control measures about looking patterns interspersed
- Repeated measures: 24 trials per session, 15 sessions over 1-2 months.
- Also collect mood survey

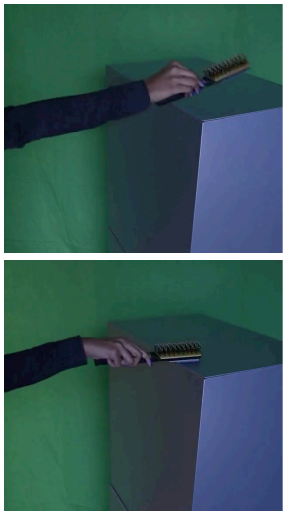
Support comparisons

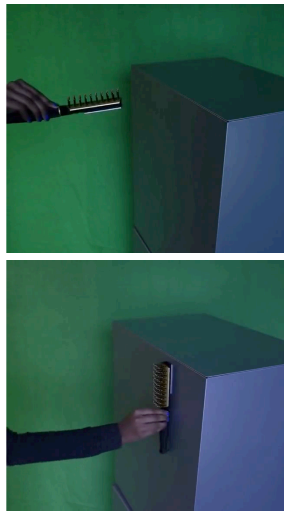
Stay





Fall





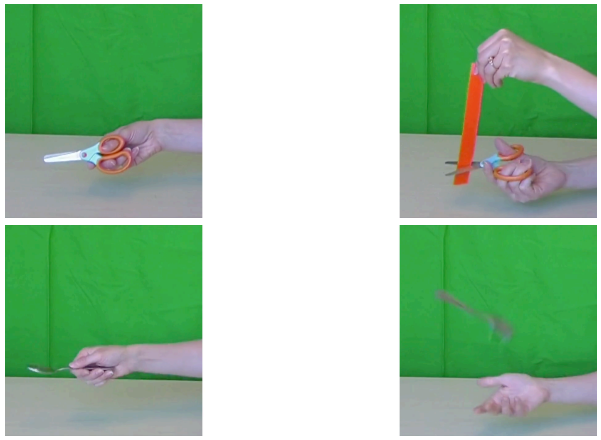
Dense measurements of intuitive physics judgments

Ongoing study: "Your baby, the physicist"

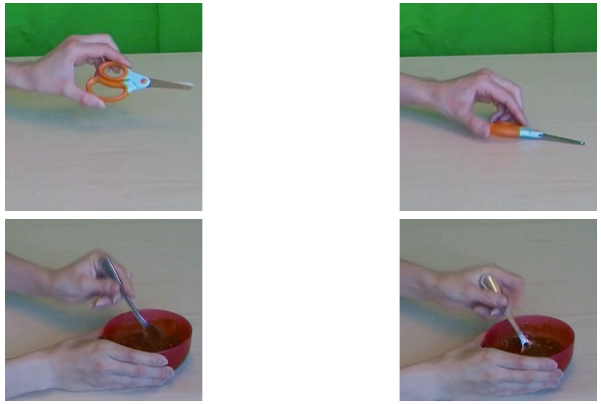
- 50 infants (4- to 12-months) participate online in a preferential looking study
- Each trial: introduce an object, then show a pair of more- and less-surprising events (loosely related to gravity, inertia, and support)
- Control measures about looking patterns interspersed
- Repeated measures: 24 trials per session, 15 sessions over 1-2 months.
- Also collect mood survey

Calibration

Salience



Similar

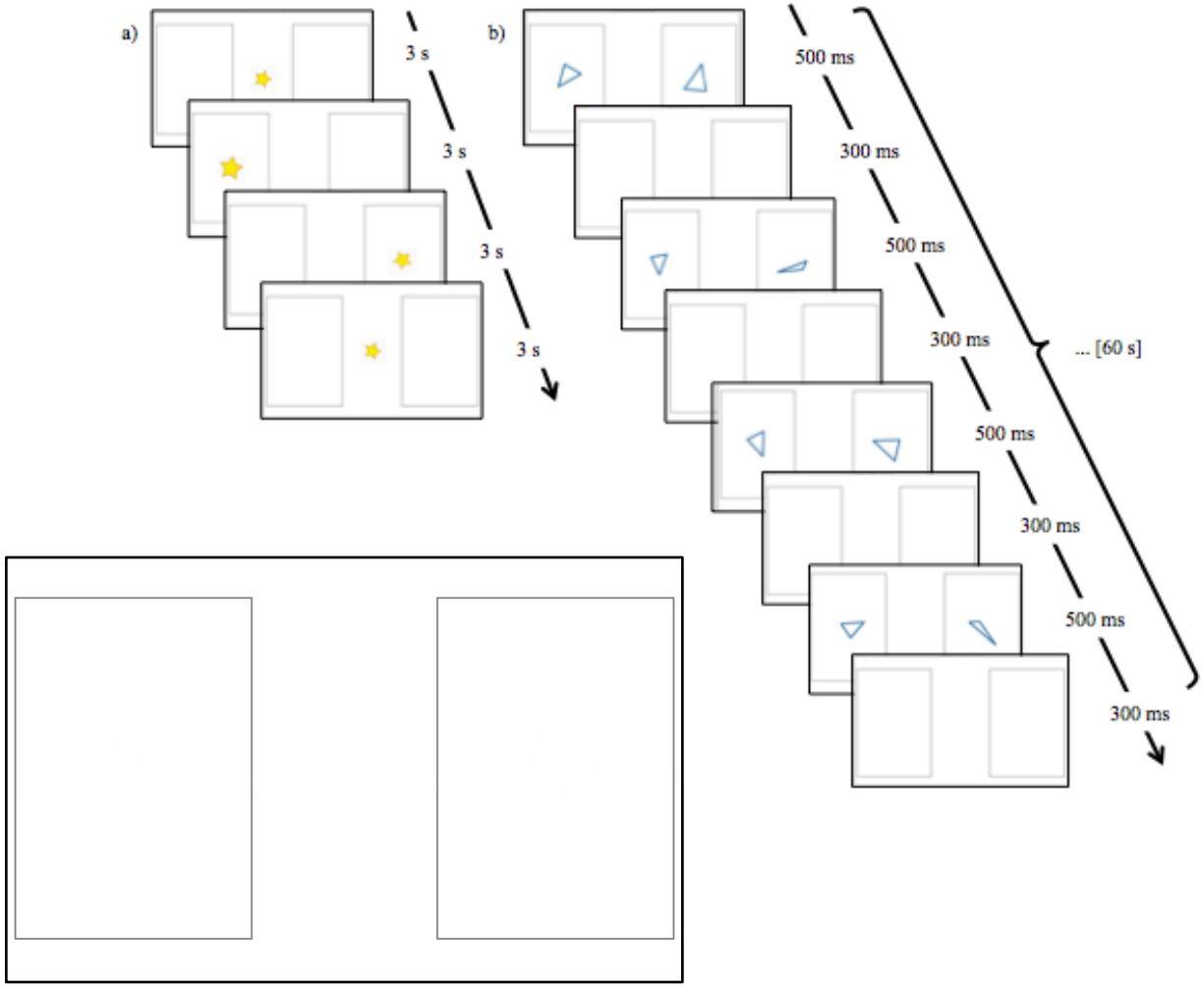


Dense measurements of intuitive physics judgments

- How stable are kids' looking patterns across sessions?
  - Contributions of age, child, session to variance on controls & test
  - How much can controls – stickiness, sensitivity – explain variation? How do these relate to measures of mood?
- Overall structure of tasks
  - Gravity, inertia, support: do responses cluster by "concept"?

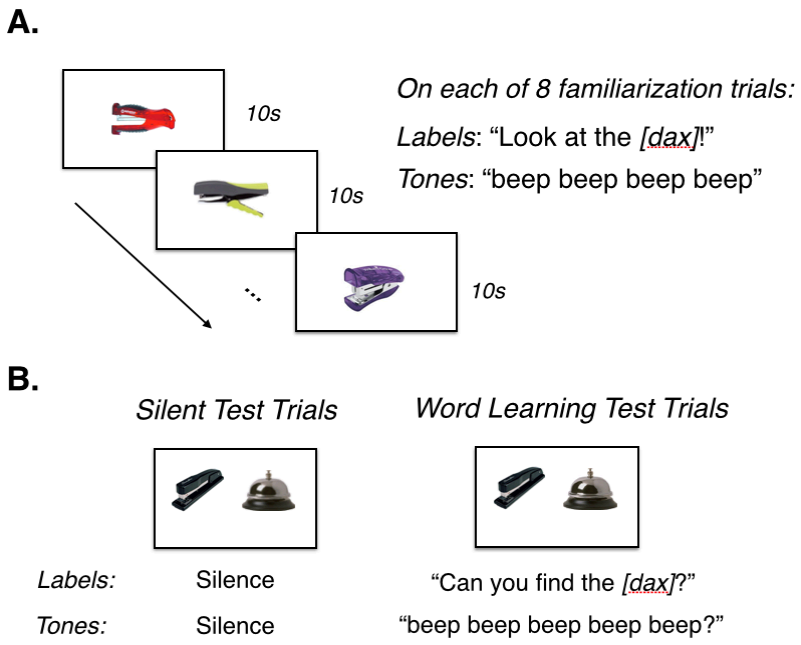
Shape perception (with Molly Dillon, Liz Spelke)

- Are infants sensitive to shape changes within triangles?
- Two triangle "streams": alternating in size only, or in shape and size
- 4 trials, each starting with calibration
- Custom frame for trial – generalizable example of "state changes" and on-the-fly graphics



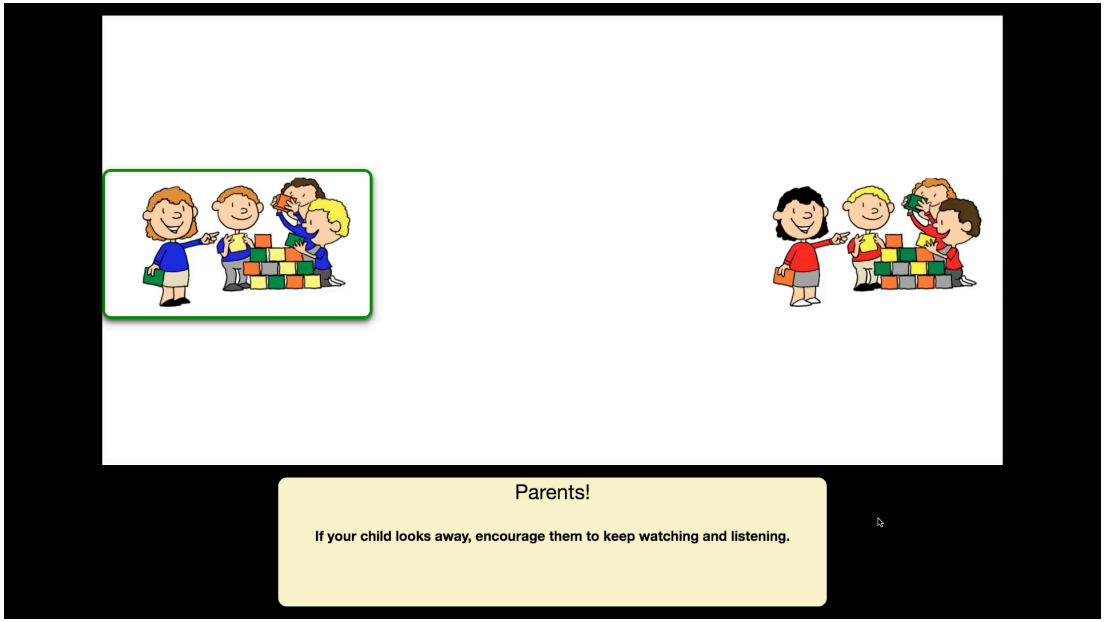
Facilitation of visual category formation by labels (with Bria Long, Mike Frank)

- Do linguistic labels help infants represent visual categories?
- Familiarization, silent test, and word learning trials supported by a new looking time frame
- Motivated development of a general-purpose randomizer



Ingroup/outgroup moral obligations (with Lisa Chalik, Yarrow Dunham)

- How do preschoolers view moral obligations within vs. across social groups?
- Custom "storybook" frame supports introduction segments, two-alternative forced-choice questions



Evaluating politeness of utterances (with Erica Yoon, Mike Frank)

- What signals do preschoolers use to judge politeness?
- Custom "dialogue" frame supports showing characters + what they say, two-alternative forced-choice questions

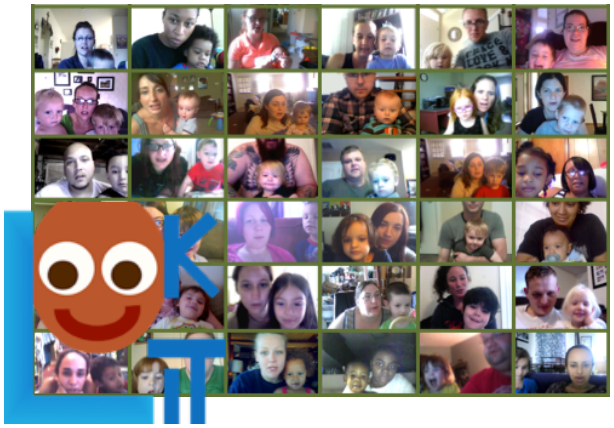
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit

Lookit: a platform for everyone

A vision for Lookit:

- Large collaborative "online lab" run by group at MIT
- Single participant interface; shared participant pool, servers, recruitment efforts
- Training, IRB coordination, study implementation, design support
- Researchers independent, but centralized approval of studies
- Support/incentives for best practices

Lookit: a platform for everyone

Mission: Lower barriers to conducting and participating in rigorous, reproducible developmental research that advances our understanding of development or our children's quality of life.

We are committed to…

- Open source development
- Encouraging data and protocol sharing
- Encouraging best practices in experimental design
- Advancing our understanding of methods
- Recruiting a representative participant pool
- Respecting participants' time and parents as partners in discovery
- Enabling non-traditional developmental researchers and supporting work that benefits children

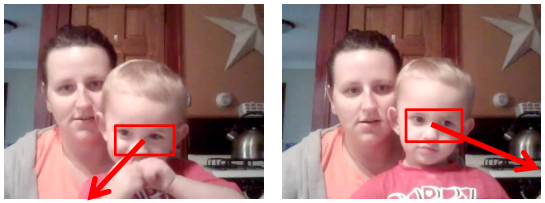
What's left to figure out

Recruitment

Additional functionality

	est. dev. weeks
Lookit website:	
Highlight new feedback on Lookit	1
Sort/display available studies by child's eligibility	1
Experimenter:	
Expanded set of useful, generalizable frames	6
Ability to export a static version of an experiment	4
GUI for consent coding (or centralize)	4
Speed up preview when only JSON changes	2
Support for payment to participants	4
"Study slot" cap and management by admin	3
Store JSON schema text directly	1
More powerful counterbalancing possibilities	3
Tool to upload feedback	2
Expanded email functionality	4
General access to API	1
Improve CSV data provided to researchers	1
Download arbitrary list of videos	2
Allow families to see their videos	3
Automatic concatenation, annotation of videos	6
Video recording:	
Switch from Flash to HTML5 video	8
Investigate missing audio issue	2
Salient way to check whether mic is working	1
Supporting online processing of video frames	4
TOTAL	**63**

Automated gaze coding



- Relax coding bottleneck
- Online contingent displays

many thanks to

participating children and families on Lookit and at the Boston Children's Museum!

Laura Schulz
Rebecca Saxe
Liz Spelke
Josh Tenenbaum

Cynthia Fisher, Kathleen Corriveau, and Ernő Téglás who graciously shared original stimuli and/or data

Kris Brewer; Melissa Kline, and 9.S93 students who made stimuli: Emily Lydic, Larissa Pachuta, Tina Zheng, Alice Lu; Max Siegel

David Szakaly: http://dvdp.tumblr.com

Research assistants: Susie (Sol Jol) Lee, Cindy Zho, Alice Lu, Jessica Wass, Chloe Joray, Vivienne Wang, Shirin Shivaei, (Nia) Dasul Jin, Jean Chow, Jasmine Gums, Scout Brisson, Daniela Carrasco, Jean Yu, DingRan (Annie) Dai, Joseph Alvarez, Junyi Chu, Rianna Shah, Katy Hanling, Hope Fuller-Becker, Audrey Ricks, Jessica Zhu, and Alice Wang.

Lookit beta testers: Molly Dillon, Liz Spelke, Mike Frank, Erica Yoon, Bria Long, Yarrow Dunham, and Lisa Chalik

CENTER FOR Brains Minds+ Machines

NSF Grant 1429216
NSF Graduate Research Fellowship Program (Grant No. 1122374)
Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216
American Association of University Women

Age 1

Me: No pens in the eyes, Remy.
Remy: Put it in the nose? Put it in the mouth? Put it in the Mama nose? Put it in the Mama mouth?

Age 2

I don't wanna stay in this world… my belly hurts. […] I wanna go back to the BCS, because in the BCS I'm feeling fine.

Remy: Are you alive?
Me: Yes.
Remy: AM I?
Me: Yes.
Remy: But I don't know what to do. I have NO IDEA.

Age 3

Everybody has something they're afraid of. I'm afraid of dolphins. […] What makes my belly hurt is, something might go wrong with a dolphin.

Age 4

Remy: What do you love about yourself?
Me: Uh, I love all the things I can do.
Remy: And all the things you try to do, but you can't do? And all the things you want to do, but you can't do?

Remy: If they go together, they don't fight. If they go apart, they fight. So if the two countries want to go towards each other...
Me: Countries actually don't move.
Remy: PLATE TECTONICS. Do you know about plate tectonics?

Age 5

Rough days are just a thing you have to go through. When you have a rough day I'll just be extra kind to you.

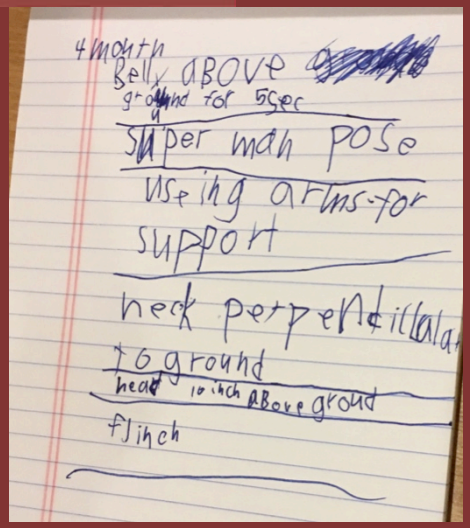
Age 6

You're my favorite mama… even though there are others that are better… which you know.









4month
Belly ABOVE
ground for 5sec
Super man pose
useing arms for
support
neck perpendicular
to ground
head 10 inch ABove ground
flinch

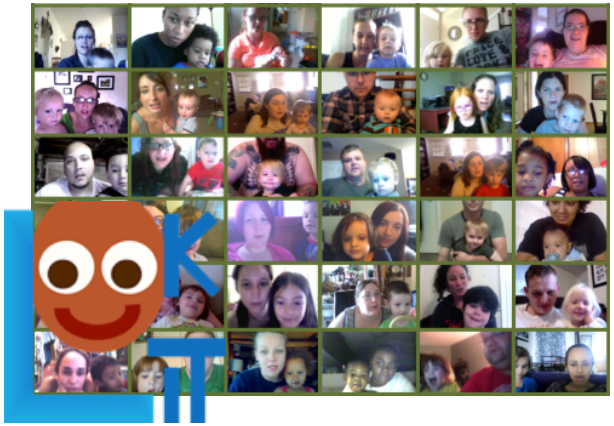
Outline

Why move online?

- Practical challenges in testing kids
- Illustration: looking for split-brain babies
- Advantages of online testing

Adapting to the online environment

- Test study protocols & results
- Demographics
- Looking measures
- Verbal responses
- Yield



Current Lookit platform & next steps

- How it works
- Example use cases
- A vision for Lookit