

Supplementary Material

An asynchronous, hands-off workflow for looking time experiments with infants

Table of Contents

Experimental design for stimulus selection Zoom experiments	1
Efficiency	2
Goal	2
Physics	3
Regression tables for stimulus selection linear mixed effects models.....	3
Supplementary results for Zoom (stimulus selection) experiments	5
Regression tables for linear mixed effects models on in-lab and hands-off experiments	5
Supplementary results for validation analysis of hands-off workflow	6
Exploratory regression analyses for effects of age and sex	6

Experimental design for stimulus selection Zoom experiments

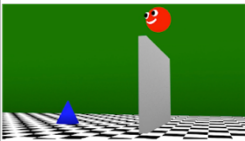
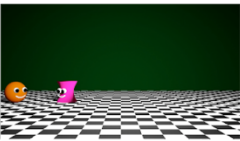
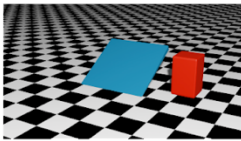
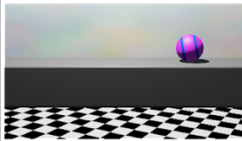
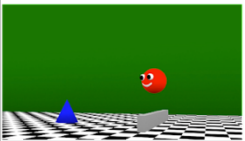
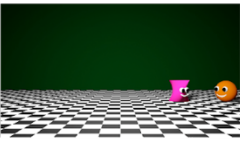
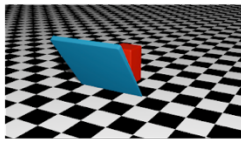
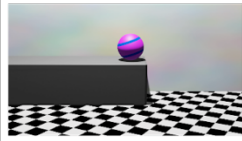
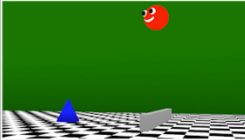
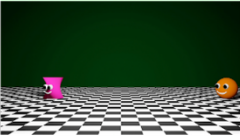
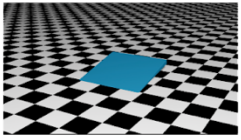
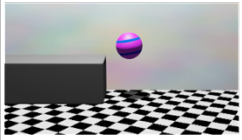
	Efficiency	Goal	Solidity	Support
Familiarization				
Expected outcome				
Unexpected outcome				
Description of unexpected outcome	<i>Agent makes high, inefficient jump over new obstacle</i>	<i>Pink agent moves away from orange character</i>	<i>Fan rotates through object</i>	<i>Ball floats in mid-air</i>

Figure S1: All four Zoom experiments include familiarization trials, two test trials with expected outcomes, and two test trials with unexpected outcomes.

Efficiency

Described in Chuey et al. (2021).

Goal

The goal study uses a violation of expectation paradigm with six familiarization trials and two pairs of test trials. In the familiarization trials, infants saw one pink actor at the center of the screen and one orange target character off to one side. The target jumps up and down, and the actor turns to look and then moves to the target. In the test trials, infants saw the same pink actor at the center of the screen, but for the first time, the target was at the opposite side of the screen. Again, the target jumps up and down, and the actor turns and then moves. In the expected test trial, the actor turns and moves towards the target, pursuing the same goal as that in the familiarization trials via a novel path. In the unexpected test trial, the actor turns and moves away from the target, moving on the same path as that in the familiarization trials.

All trials were preceded by an attention-getting animation of a star. Each trial would stop after 60 seconds had elapsed or until infants looked away for 2 consecutive seconds. The minimum looking time per trial (non-consecutive) is 3 seconds; the trial would be automatically rerun if infants did not accumulate 3 seconds of looking before they looked away. During all trials, stimuli played on a loop for the entire trial, with interleaved black screens. The experiment was also bookended by two calibration events. The calibration event before the fam trials attracted the infant's attention to the center, top, bottom, and left/right sides of the screen to guide online and offline annotation of behavior. The calibration event after the test trials was included for experimental purposes to be potentially used in subsequent work to collect better calibration images.

To counterbalance, participants saw one of two orders of test events: 1) expected, unexpected, unexpected, and expected or 2) unexpected, expected, expected, and unexpected. Whether the target character was on the left or right side during the familiarization trials was also counterbalanced. This resulted in four separate conditions that infants were randomly assigned to.

Physics

For the physics studies, each infant was tested using two tasks: solidity and support, modeled after classic behavioral work in infant cognition (Baillargeon et al. 1992; Needham & Baillargeon, 1993). Each of these tasks employs a violation of expectation paradigm, with six familiarization trials, and two pairs of test trials. In the familiarization trials for the solidity condition, a square, blue fan lies flat and a red cuboid falls from the top of the screen and next to the fan. Once the red cuboid lands, the blue fan moves 180 degrees toward it and then 180 degrees back to its original place. In the test trials in the solidity condition, the object falls *behind* the fan rather than next to it. In the expected test trial, after the object lands, the fan rotates about 120 degrees until it hits the cuboid, and then the fan rotates back to its original place. In the unexpected test trial, after the object lands, the fan rotates 180 degrees, covering the object. The fan then rotates 180 degrees back to its original place, revealing the object at its original location.

In the familiarization trials of the support condition, a pink ball with blue stripes rolls from the left to the right across a raised grey platform from one end of the screen to the other. In the expected test trial, the platform extends only slightly over the middle of the screen. The ball rolls along the platform, slows down, and stops at the edge. In the unexpected test trial, the platform is set up just as it is in the expected trial, except the ball rolls along the surface, reaches the edge, and continues rolling past the edge. It hovers in mid-air once it passes the edge. Like with the goals study, all trials of the physics study were preceded by an attention-getting animation of a star, and the experiment was bookended by the two calibration events.

We counterbalanced the order of test events (i.e. either expected, unexpected, unexpected, and expected or unexpected, expected, expected, and unexpected). For babies that saw both conditions, they were randomly assigned to watch the same order of test events across both sessions.

Regression tables for stimulus selection linear mixed effects models

All the linear mixed effects models included subject-wise random intercepts, as pre-registered in https://osf.io/jvqdg/?view_only=6722cddbafc94bf4a1dac7e027360217.

	β	SE	t	df	p
Intercept	2.37	0.19	12.39	89.76	<.001
Test type	0.28	0.14	2.08	66.60	.021
Trial number	-0.01	0.06	-0.17	76.26	.862

Table S1: Regression table for Zoom efficiency experiment

	β	SE	t	df	p
Intercept	2.84	0.19	15.30	75.24	<.001
Test type	0.20	0.12	1.72	56.54	.045
Trial number	-0.15	0.06	-2.72	60.69	.009

Table S2: Regression table for Zoom goal experiment

	β	SE	t	df	p
Intercept	2.08	0.16	12.97	108.59	<.001
Test type	-0.07	0.12	-0.55	80.35	.292
Trial number	0.05	0.05	1.03	79.58	.308

Table S3: Regression table for Zoom solidity experiment

	β	SE	t	df	p
Intercept	2.53	0.15	16.70	154.30	<.001
Test type	0.23	0.11	2.18	115.38	.016
Trial number	-0.16	0.05	-3.26	123.16	.001

Table S4: Regression table for the Zoom support experiment

Supplementary results for Zoom (stimulus selection) experiments

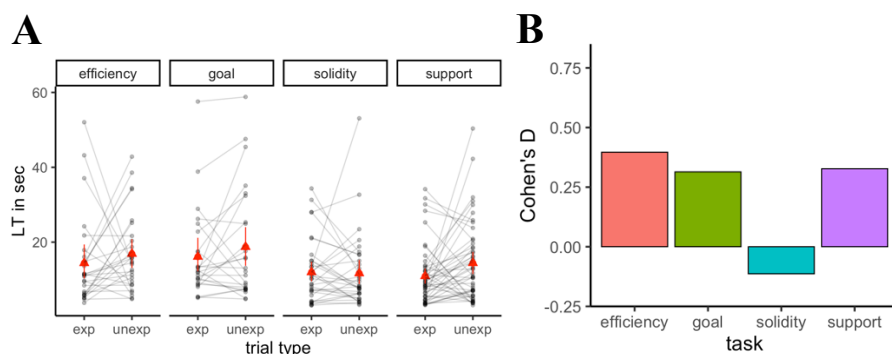


Figure S2: Supplementary results for stimulus selection experiments: (A) Raw looking time data and (B) Cohen's D on Zoom VoE experiments

Regression tables for linear mixed effects models on in-lab and hands-off experiments

The table for the Zoom workflow discussed in the 'Testing the hands-off workflow' section is shown in Table S1.

	β	SE	t	df	p
Intercept	2.66	0.18	14.63	43.28	<.001
Test type	0.30	0.10	3.06	89.83	.003
Trial number	-0.14	0.03	-4.90	92.65	<.001

Table S5: Regression table for the in-lab results

	β	SE	t	df	p
Intercept	2.41	0.10	23.56	350.65	<.001
Test type	0.21	0.07	3.29	291.70	.001
Trial number	-0.01	0.03	-0.42	296.71	.677

Table S6: Regression table for the hands-off workflow results

Supplementary results for validation analysis of hands-off workflow

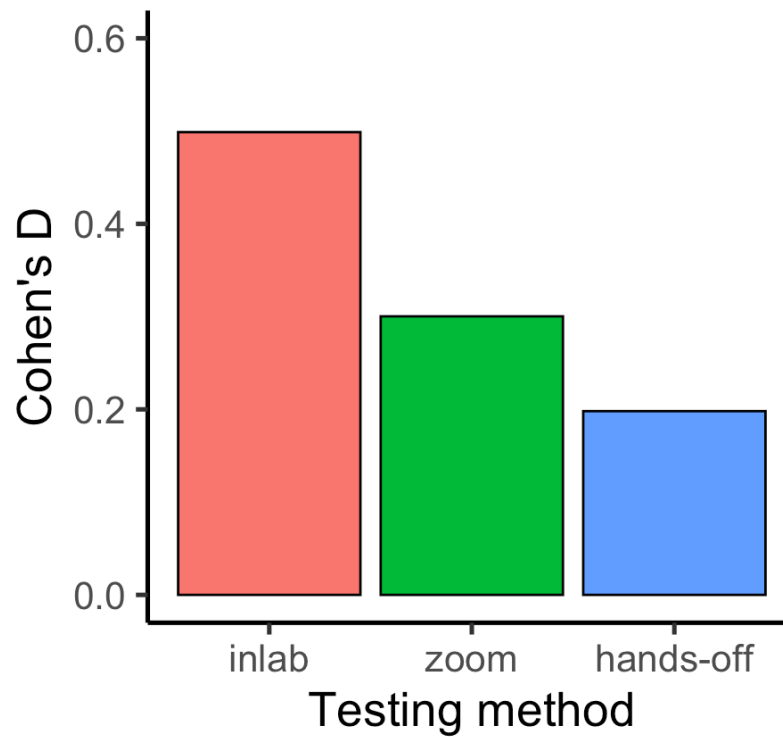


Figure S3: Cohen's D on VoE experiments run in-lab, on Zoom and hands-off

Exploratory regression analyses for effects of age and sex

	β	SE	t	df	p
Intercept	2.13	0.29	7.45	171.85	<.001
Test type	0.24	0.05	4.68	443.24	<.001
Age	0.00	0.00	1.35	149.13	.178
Trial number	-0.06	0.02	-2.76	467.16	.006

Table S7: Regression table for testing for main effect of age.

	β	SE	t	df	p
Intercept	2.55	0.09	27.18	337.71	<.001
Test type	0.24	0.05	4.67	443.35	<.001

	β	SE	t	df	p
Sex	-0.09	0.11	-0.85	150.43	.398
Trial number	-0.06	0.02	-2.89	476.10	.004

Table S8: Regression table for testing for main effect of sex.

	β	SE	t	df	p
Intercept	2.05	0.31	6.53	241.60	<.001
Test type	0.41	0.26	1.59	441.09	.056
Age	0.00	0.00	1.51	221.69	.133
Trial number	-0.06	0.02	-2.74	466.20	.006
Test type:Age	0.00	0.00	-0.66	441.53	.510

Table S9: Regression table for testing for interaction between test type and age.

	β	SE	t	df	p
Intercept	2.55	0.10	26.17	374.56	<.001
Test type	0.24	0.07	3.29	442.55	.001
Age	-0.09	0.12	-0.74	229.31	.457
Trial number	-0.06	0.02	-2.88	475.19	.004
Test type:Sex	0.00	0.10	-0.03	442.42	.974

Table S10: Regression table for testing for interaction between test type and sex.

	β	SE	t	df	p
Intercept	1.99	0.46	4.33	153.87	<.001
Testing Method (in-lab)	-1.31	3.83	-0.34	125.02	.732
Testing Method (Zoom)	0.74	1.09	0.68	147.16	.496

	β	SE	t	df	p
Age	0.00	0.00	1.39	149.66	.166
Trial number	-0.06	0.02	-2.74	460.54	.006
Testing Method (in-lab):Age	0.01	0.02	0.38	124.98	.705
Testing Method (Zoom):Age	0.00	0.00	-0.72	147.78	.472

Table S11: Regression table for testing for interaction between testing method and age.

	β	SE	t	df	p
Intercept	2.70	0.10	25.86	246.00	<.001
Testing Method (in-lab)	-0.22	0.23	-0.98	133.92	.327
Testing Method (Zoom)	-0.02	0.20	-0.08	160.65	.936
Sex	-0.15	0.13	-1.17	148.35	.245
Trial number	-0.06	0.02	-2.76	459.36	.006
Testing Method (in-lab):Sex	0.30	0.31	0.96	128.35	.337
Testing Method (Zoom):Sex	0.14	0.30	0.44	153.91	.658

Table S12: Regression table for testing for interaction between testing method and sex.