Supplemental information

Methods

Data and analysis code: Please see the Open Science Framework repository https://osf.io/9b4qj/.

Recruitment: Participants were primarily recruited via Amazon's Mechanical Turk (AMT) service and were paid for participation, regardless of whether they completed the entire study. Parents verified their participation by providing a code shown during the study or by providing a brief physical description to allow us to find a matching video, and by answering a short question about what happened during the study. We limited repeat submissions through the AMT interface using Unique Turker (Ott, 2013); parents who wanted to complete the same study with multiple children in the age range emailed us for instructions to submit a separate HIT. Submissions that did not correspond to actual participation or attempted participation were very rare and lacked answers to the verification questions.

To avoid requiring parents to provide an email address to complete the HIT, which would violate the AMT terms of service (Amazon, 2014), we encouraged parents uncomfortable with providing an email address to use a fake email (e.g., [AMTID]@mit.edu) for registration. Because external registration and video data could be used by a malicious Requester for marketing, spamming, or harassment, we advise caution in advertising studies on AMT: requirements to share personal data should be minimized and potential uses made clear. Due to ambiguity about whether external registration and video collection are allowed in cases where they are not used for any objectionable purposes, we do not view AMT as a definite long-term recruitment platform. However, participants who reviewed our tasks on TurkOpticon (https://turkopticon.ucsd.edu), a service that collects reviews from AMT workers so that they can protect against exploitative HITs, gave unanimously positive reviews that thoughtfully addressed potential concerns.

Birthdates:
Just as in the lab, we relied on parents to provide their child's date of birth in order to determine age and study eligibility. We asked parents the child's birthdate both at registration and at the end of each study. A pop-up window appeared informing parents if the participating child was outside the age range or had previously completed the study, but parents were permitted to proceed in all cases. Parents were encouraged to list a different date within the same week if uncomfortable providing an exact birthdate. 90% of participants provided exactly the same birthdate at registration and at the end of the study. In cases of disagreement, birthdates were averaged or corrected when possible based on the video. There was a difference of under one month 6% of the time; in these cases we used the average birthdate. Differences of up to one year occurred 2% of the time; to be conservative, we did not include these children unless both birthdates were in the age range or the correct date was obvious from the video (e.g., because one provided date was in the present year). There were four differences of over a year (0.9%); in three the cases this was due to the parent entering the same day in different years and in all cases it was apparent from the video which was the correct birthdate. One child in Study 1 was excluded for being older than 18 months based solely on video, as she appeared to be at least 3 years old.

Server and study setup: These Lookit study protocols are written from scratch in JavaScript using jQuery; researchers specify periods during the study to record using a custom JavaScript library which interfaces with Flash. The studies described here use 8-13 "clips," or periods of recording, per session. Each clip is streamed to Amazon S3 cloud storage using Wowza Media Server. Video quality is limited by the user's upload speed; we chose not to store video locally so that studies could be completed in a web browser. The server hosting the studies establishes a secure https connection to transfer all data. Account, demographic, and session data are each stored in a MongoDB collection on the server. Videos are automatically downloaded daily to a computer in the lab; for each study, a spreadsheet of all sessions is generated based on the session database. For links to current Lookit code, please see https://osf.io/9b4qj/.

We distinguish between users (adult account holders on the Lookit site), participants (children who begin at least one study), and sessions (records of a participant completing a particular study). Multiple participants may be associated with the same user and multiple sessions with the same participant.

Coding for fussiness, distraction, and parent interaction: In the studies using looking measures (Studies 1 and 2), individual clips were coded for several events including child fussiness, distraction, and parent interaction events, defined as shown in Table S1. Two blind coders recorded which events occurred during each clip. For further details on coding procedures see the coding manual.

Event	Definition	Coded for
Fussy	Child cries or child squirms in an attempt to physically face away from the screen and/or get off the parent's lap	All clips
Distracted	Any lookaway from the screen clearly caused by an external event	
Parent talks	Parent talks to child in a way that potentially affects looking behavior	Clips where the parent's eyes should be closed
Eyes open	Parent's eyes open for most of trial	
Peek	Parent's eyes are open briefly	
Eyes open at start	Parent's eyes are open for up to the first few seconds of the clip, to see that playback starts on the screen	
No eyes	Parent's eyes are not visible on webcam	
Point	Parent points at the screen	

Table S1: Qualitative events coded in Studies 1 and 2.

Parent cooperation

Parent response (N = 250)	Number of parents	Parents with eyes open during at least one test trial, based on video coding	Parents peeking during at least one test trial, based on video coding
Yes	154 (65%)	3	26
Tried to/peeked	60 (25%)	8	31
No	23 (10%)	11	5
No response	13 (5%)	1	4

Table S2. Coded parent responses to the question "Did you always keep your eyes closed?" at the end of the study. All coded video from Studies 1 (N = 112) and 2 (N = 138) included; there were four test trials in Study 1 and three in Study 2.

In order to investigate parents' compliance with requests to close their eyes during portions of the study, parents were asked at the conclusion of Studies 1 and 2 whether they actually kept their eyes closed when requested. Their responses were coded as "yes," "no," or "tried to/peeked." A summary of the responses from all potentially usable records, along with coded behavior during test trials, is shown in Table S2. Parents were asked to close their eyes at the beginning of each of six 20 s trials in Study 1; the last four were considered test trials. In Study 2, parents were asked to close their eyes for clips 4-6 and 11-13, without breaks in between. The last three 8 s clips were considered test trials. Reasons given for incomplete compliance included needing to care for another child or respond to or supervise the child participating (n = 15), being unsure about the directions or forgetting on the first trial (n = 6), curiosity (n = 3), wanting to check that videos shown were appropriate (n = 6), technical problems (n = 4), and believing it was unnecessary (n = 2). Parents who said that they did always keep their eyes closed when requested ("Yes" response) were indeed more compliant during test trials: chi-square tests of independence revealed relationships between "Yes" responses and both having eyes open during at least one test trial ($p < 0.01$) and peeking during at least one test trial ($p < 0.01$).

Recoding looking measures: In cases of disagreement between coders (see main text for study-specific criteria) regarding looking time or preferential looking measures, individual clips were recoded, again by two blind coders. Coders asked to recode might or might not overlap with the original coders. Each coder was asked to check one of the original coding records but did not know in which direction this coding differed from the other original coder's. The function of recoding was to avoid reduce obvious human error as a source of disagreement.

Figure legends

Figure S1: Median absolute difference between looking time measurements calculated based on downsampled coding vs. at original framerate (29.97 Hz). Error bars show IQR. We started with blind coding from original video of 283 infants and toddlers tested in a quiet room at the Boston Children's Museum on a total of 2097 looking time trials. For the 1997 trials with a continuous lookaway of at least 1 s, we downsampled the coding by rounding the starts and ends of recorded looks to the nearest simulated frame start time (to the nearest ms), then recomputed looking time based on the downsampled coding.

References

Amazon (2014, Dec 2). Amazon Mechanical Turk Participation Agreement. Retrieved from https://requester.mturk.com/policies/conditionsofuse

Ott, M. (2013). Unique Turker. Retrieved from https://uniqueturker.myleott.com/

Video legends

Note: All videos are from sessions where parents chose a privacy level of "free" and confirmed that publicity and educational use of the video are acceptable.

SI Video 1: A collage of six examples of informed consent recordings made by parents before beginning studies on Lookit.

SI Video 2: An example of a high-framerate video collected on Lookit. The effective framerate is estimated at 13.01 frames per second (see main text Methods). This clip is from a calibration trial in Study 2.

SI Video 3: An example of an excluded low-framerate video collected on Lookit, collected during a calibration trial in Study 2. The effective framerate is estimated at 1.47 frames per second (see main text Methods); videos with framerates under 2 frames per second were not used in the analysis.