Using Tablets to Collect Data from Young Children

Michael C. Frank, Elise Sugarman, Alexandra Horowitz,

Molly L. Lewis, & Daniel Yurovsky

Department of Psychology, Stanford University

Abstract

Mobile, touch-screen devices are increasingly ubiquitous in children's lives. The extensive use of such devices provides both a challenge and an opportunity for developmental psychologists. While more research is needed to understand the nature and consequences of young children's interaction with such devices, they nevertheless present an exciting opportunity for data collection. We describe a simple method for creating cross-platform, interactive tablet experiments using open web-based resources. We illustrate this method by collecting data from children 1–5 years old in a simple word-recognition paradigm and show that it yields reliable accuracy and reaction time data. Tablets should hence be considered by researchers as a viable method for collecting low-cost, well-controlled developmental data.

Introduction

Since Apple debuted the iPad in 2010, the ubiquity of personal tablets—interactive, touchscreen-based computing devices, typically larger than a phone but smaller than a laptop—has skyrocketed. By 2013, an estimated 75% of families with young children in the United States owned a mobile device, with 40% of households owning tablets (20% of lower-income families and 63% of higher-income families; Rideout, 2014); tablets were estimated to constitute 19% of children's overall screen time from ages 0–8. Accordingly, the market for downloadable applications targeted to children has increased dramatically, representing 72% of the top selling apps in the iTunes store, and over 80% in the Education category (Shuler, Levine, & Ree, 2012), although one report suggests the viewing photos is the most popular activity for children under age 4 (Cristia & Seidl, in press). This explosion in popularity has created a surge of interest in both the consequences of tablet use for children and the potential for using these devices as scientific tools.

Nevertheless, there are significant challenges involved in using tablets for research. These include the difficulty and potential costliness of developing custom applications ("apps") for individual experiments, technical problems with app distribution and cross-device compatibility, and the lack of data confirming the reliability of data collected with tablets. Perhaps for these reasons, there is currently a paucity of published articles that use tablets to collect developmental data.

The goal of the current article is to address these challenges. We begin by briefly reviewing research on the use of tablets in children. We then describe a method for using standard, freely-available web-development tools to create tablet-based experiments. Although this method requires some programming experience, it is significantly easier than developing tablet-native applications and has relatively few drawbacks from the research perspective. We next show how this method can be used to implement a simple

word-recognition experiment. Our data suggest that the tablet platform is an engaging method for data collection that yields reliable accuracy and reaction time from children two years old and perhaps even younger.

Prior Work on Children's Tablet Use

The primary focus of research on tablets with children has been evaluation of their educational potential. The starting point for these investigations is the suggestion that children learn less from videos than they do from live demonstrations, a phenomenon known as the video deficit effect (e.g. Anderson & Pempek, 2005; DeLoache et al., 2010; Kuhl, Tsao, & Liu, 2003). These findings influenced the American Academy of Pediatrics (AAP) to uphold a policy recommending no "screen time" for children under age 2 (Brown, 2011).

The experiences offered by touch-screen devices may be distinct from passive viewing of television and computer screens. Christakis (2014) notes that the AAP's published media recommendation of no screen time was made before the release of tablet devices, and suggests that the policy should be revised to distinguish and accommodate some time with touch-screens. He proposes that tablets can offer a number of potential benefits that many traditional toys do not. Specifically, tablets are reactive (responding contingently to the childs actions), interactive (eliciting responses from the child based on his or her actions), tailorable (modifiable according to age or preference), progressive (can increase in difficulty or complexity with experience), highly portable, and able to promote joint attention (a child and caregiver can interact with the program together).[1]

---

[1]Indeed, a joint statement by the National Association for the Education of Young Children and the Fred Rogers Center for Early Learning and Children's Media suggests that a single "screen time" recommendation oversimplifies the possible benefits of new technologies, and educators should explore the potential pros and cons of interactive medias to determine when and whether they are appropriate to support children's learning (Radich, 2013).

While screens are not recommended as replacements for live interactions, interactive technologies may pose a number of benefits over other passively-viewed screen alternatives. Some work suggests that practice with interactive games may benefit attention and motor control (Bavelier, Green, & Dye, 2010), and more generally screen time with tablets can promote independence, opportunities for physical closeness, and feelings of fun and accomplishment Rvachew (2014). The contingent feedback and individualization offered by mobile devices may thus make them a better medium for learning and engagement than traditional view-only screens. Such benefits may also make them valuable tools for research and evaluation.

In addition, young children do demonstrate some learning from interactive screen content, suggesting that "video deficit" effects may be limited to—or at least strongest in—passive viewing of pre-recorded material. In one study, Roseberry, Hirsh-Pasek, and Golinkoff (2014) found that 24- to 30-month-olds successfully learned a novel verb from a live interaction via video chat, while a yoked group who saw the same input with non-contingent presentation failed to learn. Young children also show some evidence of transferring learning between screens and physical objects. Zack, Barr, Gerhardstein, Dickerson, and Meltzoff (2009) examined 15-month-olds' ability to transfer knowledge learned via a touch-screen or physical toy to another example either in the same presentation modality or across modalities. Children were most successful when toys were presented in the same modality, but children in the cross-modality conditions performed better than a no-demonstration baseline. Even before age 2, children thus show evidence of comprehending actions and goals presented on touch-screen devices and can transfer this knowledge between physical and 2-dimensional presentations.

One concern with children's performance is the accuracy of their motor control. Couse and Chen (2010) found that young children could quickly learn to provide precise motor input on tablets with minimal training using a stylus pen. They had teachers rate

the quality and complexity of 3- to 6-year-old's self-portraits created digitally on a tablet art program compared with their drawing using traditional writing tools on paper. Children were highly engaged with tablet program, their drawings were rated as typical or exceeding their typical quality of drawing, and three quarters of children on the tablet task exhibited no to little frustration during the task despite a number of technical challenges. Additionally, although researchers were prepared to give children up to four warm-up sessions using a stylus pen on tablet computer, most children advanced after a single warm-up phase. Altogether, these studies suggest that preschoolers and kindergarteners are highly motivated and competent at interacting with touch-screen tablets.

And although children show interest and remarkable capabilities to use touch-screens even before they learn to talk, some affordances of mobile tablets are easier to master from young ages than others. Aziz, Sin, Batmaz, Stone, and Chung (2014) tested 2- to 5-year-olds in Malaysia and the United Kingdom on their ability to perform the seven most common touch gestures of iPads: tap, drag/slide, free rotate, drag and drop, pinch, spread, and flick. By age 2, children mastered the tap and drag/slide gestures, and by age 3, they could reliably produce all by the spread gesture. By age 4, all children performed all of the gestures. Cristia and Seidl (in press) additionally suggest that enabling multi-touch responding–allowing touches from more than one location to be recorded simultaneously–may reduce motor control issues of children trying to interact with the tablet while holding or resting their hand on other parts of the screen.

In sum, research on tablets suggests that interactive media may provide more learning opportunities than passive viewing. Tablets are extremely engaging to children and may provide a valuable source of information, but designers of tablet materials for young children should take care to consider appropriate content and gesture capabilities for their target age group.

Why Collect Developmental Data With Tablets?

Researchers working with young children have access to a wide variety of experimental paradigms, ranging from behavioral forced-choice to eye-tracking and looking-time based measures (Aslin, 2007; Gredebäck, Johnson, & von Hofsten, 2009). Nevertheless, each of these methods has its drawbacks: behavioral methods are flexible and easy to implement, but they can be subject to experimenter biases, and often require labor-intensive offline coding. In contrast, corneal-reflection eye-tracking methods are precise and unbiased, allowing tight experimental control, but can be expensive and are difficult to implement outside of tightly-controlled lab settings. In addition, eye-tracking paradigms are typically very limited in length as they tend to be non-interactive (cf. Deligianni, Senju, Gergely, & Csibra, 2011).

Tablet-based experimental paradigms can supplement both of these data collection modalities by allowing researchers to create engaging, interactive paradigms that are low cost but nevertheless offer some of the precision and experimental control of eye-tracking paradigms. In the sections below, we offer five reasons why tablets may be valuable tools for developmental research.

Diverse measures. Tablets allow the collection of many different dependent variables, including forced-choice responding and gesture trajectories. We especially highlight the ability to measure reaction times using tablets. Reaction times provide a continuous, informative measure of cognitive process and have proven invaluable in both developmental (e.g. Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998; Marchman & Fernald, 2008) and adult (e.g. Sternberg, 1969 et seq.) studies. Behavioral paradigms that afford reaction time measurement for young children typically require time-consuming offline coding. In contrast, eye-tracking paradigms yield reaction times, but with the tradeoff that the paradigms typically have limited interactivity.

Engagement. Children across ages enjoy tablets and are highly motivated to explore new tablet-based tasks, allowing researchers to collect more data from individual children. As documented by Couse and Chen (2010) and many informal observations of infants' and young children's interest, the use of tablet-based devices may lead children to view tablet-based experiments as similar to engaging commercial apps. In addition, behavioral paradigms with young children often require a "warm-up" period so that the child is comfortable interacting with an unfamiliar experimenter; such a period can be shorter or absent to the extent that a tablet-based paradigm requires less face-to-face interaction.

Experimental control. Studies that involve interaction with an experimenter are the de-facto standard in cognitive development research, but such studies have as major drawbacks that it may be difficult or impossible to ensure that experimenters are blind to condition and hypothesis. Often hypotheses are simple enough that even "naïve" research assistants can infer the experimental structure or the desired response. While there are methods for avoiding such issues, the possibility of eliminating experimenter bias via computerized stimulus presentation is an attractive method for situations where it is possible. And even when bias is not introduced, there are significant challenges in training a group of research assistants to implement an experimental protocol uniformly. Thus, tablet-based experiments can alleviate concerns about bias and uniformity of experimental presentation.

Inclusion of special populations. Tablets can be a valuable tool for working with children who have learning or physical disabilities (e.g. Cardon, 2012; Waddington et al., 2014; Chai, Vail, & Ayres, 2014; Bertucco & Sanger, 2013)). The lack of face-to-face interaction demands make tablets especially appealing for working with children on the autism spectrum, who may find interactions with an unfamiliar experimenter especially aversive.

Scale. The low cost and high accessibility make it possible to distribute tablets to research assistants who can collect data in a wider variety of locations (though not without concern about environmental distractions). In addition, it is in principle possible that children could contribute data to experimental studies via experiments administered by their parents or caregivers.[2] It is this latter possibility that might be truly revolutionary in terms of the scope of data collection, though there are many challenges that this sort of data collection method would need to overcome.

Despite all of these advantages, there are significant technical issues related to app creation that have slowed the adoption of tablet-based methods in developmental research. Creating tablet-native apps (that can be accessed through an Android or iOS tablet's main menu) requires substantial programming expertise—and worse yet, this expertise is not transferrable across tablet platforms. Developing a native app for iOS requires programming in Objective-C using Apple's proprietary app development tools; development for Android requires programming in a Java variant. Both languages have specific software development tools (SDKs) that allow programmers to tailor their apps for various devices.

Although there are cases where having a robust, native app might lead to better performance or adoption, the cost of developing such apps is substantial. An experiment that requires substantial media or animation, use of the tablet's camera, or large-scale data transfer is likely to perform better as a natively written app. Similarly, standardized experimental tools are more easily distributed and used via Google Play or the Apple App Store, the standard distribution methods for commercial software. To create such apps, though, either an investigator must learn a special-purpose programming language, or a programmer must be employed. This latter option is more common, but costs can easily

---

[2]For an example of this method, see `http://lookit.mit.edu`, a web-based platform for looking-time experiments that engages parents as experimental confederates.

Server hosting HTML/JavaScript/CSS   Server hosting PHP script

③

①   ②

① From a web browser on the tablet, the user visits the experiment webpage hosted on a web-based server.

② The webpage interacts with the participant via JavaScript code.

③ After the last trial, the JavaScript calls a PHP script that is also on a web-based server. This script logs the participant's data to a tabular data file.
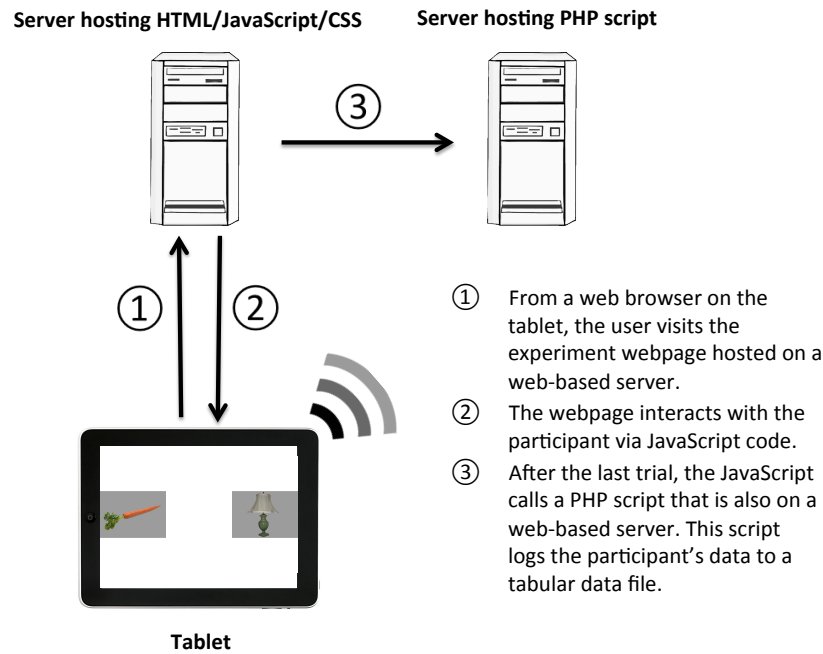
Tablet

Figure 1. A schematic of our method of collecting tablet data using a web-based experiment. Note that the two server functions can be executed by the same system; we have separated them here for clarity.

run in the tens of thousands of dollars; in this sort of scenario, making scientifically-useful changes to an app may be undesirable because of added costs.

A Web-Based Method for Developing Tablet Experiments

In contrast to costly, special-purpose methods, here we describe an alternative method for prototyping and collecting data on tablets. The method relies on creating the experiment as a simple webpage and then showing this page in the tablet's web browser. This method still requires some programming experience, but much less. In addition, it draws on open-source, freely-available tools that are relatively easy to master compared with native app development tools. Experimental "apps" made via this method are easily

shared, modified, and ported across platforms, leading to faster iteration and data collection. In addition, our method offers the ability to collect control data with adult participants on Amazon Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010; Crump, McDonnell, & Gureckis, 2013) with virtually no changes to the underlying experiment. Our method has three components (see Figure 1):

1. A JavaScript/CSS/HTML web page, which is the core of the experiment,

2. A server-side PHP script to collect data in tabular format, and

3. A tablet device running a kiosk application to present the experiment.

We discuss each below.[3]

Web-based experiment. The first component of our system is a web-based experiment created using HTML, JavaScript, and CSS. This combination is currently the most common set of tools for the creation of interactive websites. HTML (HyperText Markup Language) forms the framework for the static content shown on the website, while CSS (Cascading Style Sheets) gives a consistent set of formatting options. JavaScript then controls the interaction functionality of the website. In particular, we make use of a JavaScript library called jQuery (`http://jquery.com`) that allows HTML elements like images, layouts, text, and media to be dynamically added, subtracted, and rearranged as participants interact with the resulting page.

It is beyond the scope of this article to provide an introduction to creating JavaScript-based web experiments, but there are a wealth of available resources on this topic. These resources include both psychology-specific tutorials and more general introductions to the general HTML/JavaScript/CSS model. A sample experiment (as well as the server side code) is provided in the version control repository for this paper

---

[3]Our work has used the Apple iPad as the primary tablet platform and we provide details that are tailored for this platform. Nevertheless, and unlike other methods, our method is in principle portable to a wide range of mobile devices with only very minor modifications.

(`http://github.com/langcog/tablet_norming`) and can be viewed at
`http://langcog.stanford.edu/expts/TABLET/tablet.html`.

Hosting a web-based experiment requires server space, which is often provided by universities but can also be purchased from commercial providers. We find that the requirements on such a server are typically not high, so the costs involved are minimal.

Server-side data collection. The second component of our system is a simple script that allows the experiment to save data in tabular format. Although native apps can store information on tablets, browser-based apps such as ours cannot; thus, our method requires that the data be saved to a separate server. Every time a participant completes the experiment, their data is sent to this script, written in PHP (a simple scripting language that is often used for backend web development) and hosted on a server. The script simply appends data from the experiment to a tabular data file. The script can be hosted on any web-accessible server that has been configured to run executable files of this type. In practice, this will often be the same server that hosts the experiment itself, though it need not be. Data are then retrieved from this server periodically using SFTP (Secure File Transfer Protocol).

Tablet configuration. The third component of our system is an internet-connected tablet. The requirement for internet connectivity is perhaps the most substantial drawback of our method, relative to native apps. The tablet must either have wireless access (e.g., from the testing location or a mobile hotspot device) or else the tablet itself must have cellular connectivity. Once the tablet has access to the web, the tablet's browser can simply be directed to the experiment website.

A major challenge of presenting experiments in the tablet's web browser is to ensure that children cannot accidentally or intentionally navigate away from the experiment, exit the browser, or change perspective (e.g. by zooming). In practice, we use two tools for

this purpose: the first is Guided Access, a mode on the iPad that disables the hardware buttons. The second is Mobile Kiosk, an app that further locks the iPad into a particular view of a webpage. The combination ensures that the tablet presents the experiment as intended. With both of these systems in place, the "look" of the experiment is extremely similar to a native app.

An Experimental Test of the Reliability of Tablet Data

To test the reliability and broad developmental applicability of our tablet method, we implemented a simple word-recognition paradigm, which we tested with a cross-sectional sample of children ages 1–5 from a local children's museum. The basis of our experiment was work on the developing speed and accuracy of word recognition by Anne Fernald and colleagues (Fernald et al., 1998; Fernald, Perfors, & Marchman, 2006; Bion, Borovsky, & Fernald, 2013).

We designed a two-alternative forced choice paradigm that both tested children's recognition of familiar words and asked them to make so-called "mutual exclusivity" inferences—that a novel label refers to a novel word rather than a familiar competitor (Markman & Wachtel, 1988). While the causes of this latter inference are still a source of theoretical disagreement (Markman, Wasow, & Hansen, 2003; Diesendruck & Markson, 2001; Frank, Goodman, & Tenenbaum, 2009; Bion et al., 2013), it is nevertheless a highly replicable finding that we predicted woud show substantial developmental change across our sample.

Methods

Participants. We recruited a sample of 86 children ages 1–5 from the floor of San Jose Children's Discovery Museum. Demographics for these participants are shown in Table 1. An additional 18 children in the target age range were recruited but excluded from the sample for the following reasons: parent-reported English exposure less than our

Table 1

Demographic information for our participants.

Age Group	N	Mean Age	Proportion Male
1-year-olds	18	1.71	0.44
2-year-olds	22	2.46	0.68
3-year-olds	24	3.56	0.67
4-year-olds	22	4.51	0.59

pre-specified criterion of 75% (7), parent interference (4), issues with the iPad pre-training (4), experimenter error or technical issues (2), and self-reported developmental disability (1).

Stimuli. Visual stimuli consisted of images of sixteen familiar and eight novel objects. All were presented as cropped images on a gray background with approximately the same resolution and detail. Audio stimuli consisted of a carrier phrase ("Can you find [the target]?"), into which we spliced recordings of the target words. All recordings featured a female speaker and were created following the procedures described in Fernald, Zangl, Portillo, and Marchman (2008). We selected a variety of words from the MacArthur-Bates Communicative Development Inventory word list (Fenson et al., 1994, 2007), choosing some words that we were certain nearly all children would be very familiar with (e.g., "car," "dog," "bottle,") and others that were intended to be slightly more challenging (e.g., "monkey," "shovel," "lion"). Novel words were mono- and di-syllabic words that had been used in previous experiments of this type (e.g. "dax," "pifo," "kreeb"). We further included four fillers to break up the experiment (e.g., a picture of a train with a sentence of audio describing it).

Procedure. We created a simple page to help children learn to tap appropriately on the iPad. To advance past this training page, they needed to tap a set of five dots in random locations; each dot transformed into an "X" when tapped. After the child tapped all the dots, the experimenter could either advance the app to the experiment or rerun the training if they felt the child could use more practice.

Experimental stimuli were then presented in one of two pseudorandom lists, in which trials (which contained a pair of pictures) were counterbalanced for order, target side, and XYZ. The experiment had 28 trials in total (8 of each of the three trial types plus four fillers). The three trial types were *Familiar Word* trials, in which two familiar pictures were presented and children were asked to select the image that matched a familiar word; *ME Control* trials, in which a familiar and a novel image were presented and children were asked to select the image that matched a familiar word; and *ME Inference* trials, in which a familiar and a novel image were presented and children were asked to select the image that matched a novel word.

Results

Children were overall very enthusiastic about the experiment. One-year-olds varied in their ability to complete the study, with most completing half or two thirds but not the full 28 trials. In contrast, the other groups successfully completed the experiment at high rates. Completion rate and average number of trials completed are given in Table 2.

Accuracies for familiar word trials were above chance even for the one-year-old age group and rapidly approached ceiling performance in the two-year-olds (Figure 2). There were also no substantive differences in accuracy between Familiar Word and ME control trials. In contrast, ME Inference trials were substantially more difficult for children, with one-year-olds performing at chance, and two-year-olds showing a substantial decrement in

Table 2

Proportion children finishing the study and average number of trials out of 28 completed for all children, by age group.

Age group	Proportion finishing	Average number of trials
1-year-olds	0.44	20.4
2-year-olds	0.91	27.5
3-year-olds	1.00	28.0
4-year-olds	0.86	27.6

performance for these trials.[4] Our accuracy data thus strongly support the appropriateness of the tablet-based method for children two years old and above. In contrast, accuracy data for the one-year-olds were more mixed, even though this group was mostly older than 18 months.

We next turned to the reaction time data. Considering only reaction times for correct trials, we found a substantial skew in the reaction times we collected (Figure 3). This pattern appeared to be due to two clusters of reaction times: one larger, tighter cluster early on, representing immediate responding, and one broader slower cluster representing trials on which participants were unsure how to respond, looked to a parent,

---

[4]To test the reliability of these differences, we fit a logistic mixed effects model to the data (Jaeger, 2008); Our model specification included age group (coded as a discrete factor with one-year-olds as the intercept), trial type, and interactions of age group and trial type. Models with larger random effect structures did not converge, so we included only a random intercept for each participant. We found a significant increase in performance for all three older age groups ($\beta = 1.68$, 3.21, 2.60, all $ps < .0005$) and a significant decrement in performance for ME Inference trials ($\beta = -1.14$, $p = .0003$), but no effect in ME Control trials and no interactions between trial type and age group (all $ps > .18$).
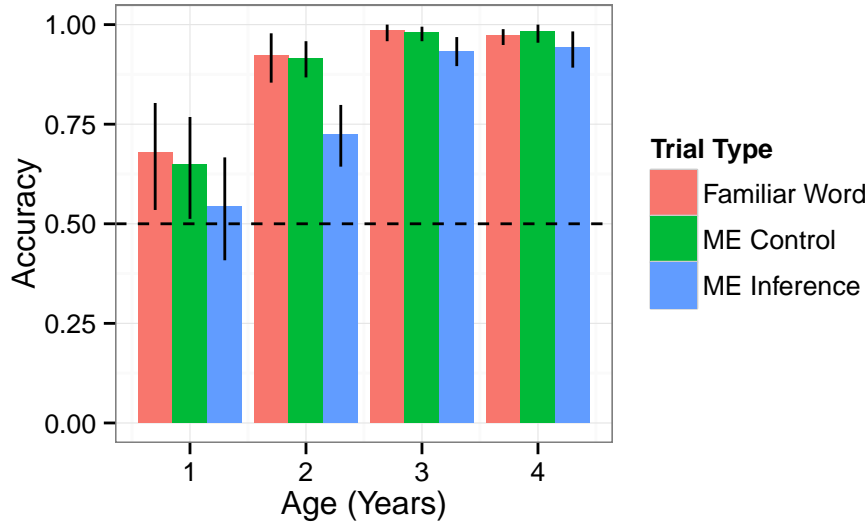
Figure 2. Accuracy in our word recognition task, plotted by age group and trial type. Dashed line shows chance performance. Error bars show 95% confidence intervals, computed by non-parametric bootstrap.

or otherwise deliberating. In an exploratory analysis, we fit a Gaussian mixture model (using the `mclust` package in `R`) and found that the best BIC values were indeed produced by a two-cluster solution, which is shown in Figure 3.

For further analysis, we excluded reaction times faster than 500 ms or slower than 4000 ms. These values were determined post-hoc, based on inspection of the distribution. They correspond well both to our intuitions about the task and to the properties of the first mixture component found by the mixture model (for which mean + 3 standard deviations was 3928 ms). Excluding based on these threshold values resulted in an exclusion of 10.9% of values.[5]

---

[5]Although this proportion seems high, consider the collection of reaction times based on eye-tracking methods. In these analyses, only half of trials are considered—those where the eyes are on the distractor item at the point of disambiguation (Fernald et al., 2008).
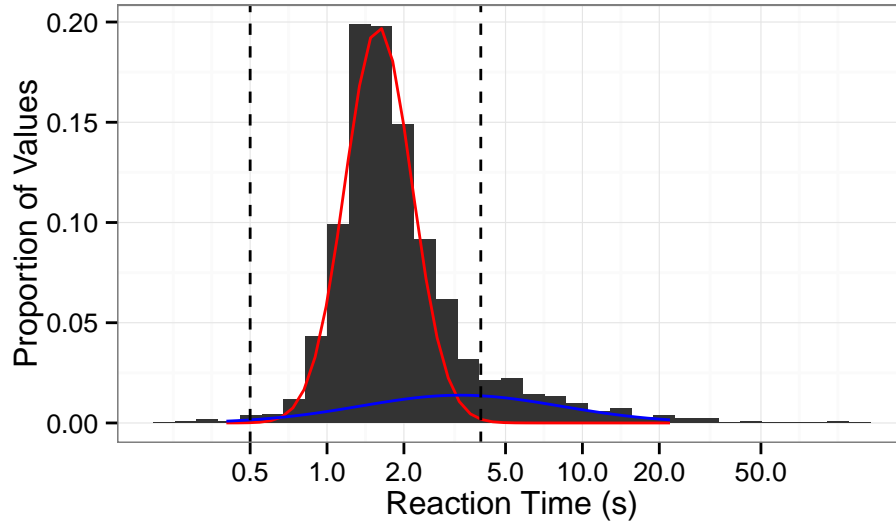
Figure 3. A histogram of reaction times across all participants. Red and blue curves show the best-fitting distributions found by a 2-component Gaussian mixture model. Dashed lines show the cutoff values we adopted (.5 and 4.0 s).

We found that older children were faster overall in our task, and ME inference trials were slower than both Familiar Word and ME Control trials for all groups except the one-year-olds (Figure 4). We again did not observe large numerical differences between Familiar Word and ME Control trials.[6] Again it appears that the iPad method yields reliable data for two-year-olds and above. Experimenting with other reaction-time outlier exclusion processes did not qualitative alter the pattern of results. Nevertheless, we recommend that investigators using this method carefully examine their distribution of

---

[6]A linear mixed effects model fit to log-transformed reaction-time data showed a significant effect of trial number such that reaction times decrease over the course of the experiment ($\beta = .001$, $p = .04$), and three- and four-year-olds were significantly faster than one-year-olds ($\beta = -.19$, .26, $ps < .005$). There were no main effects of trial type, but there were reliable slowdowns for ME Inference trials across the two-, three-, and four-year-olds ($\beta = .25$, .26, .29, all $ps < .0001$). ME control trials were also reliably faster for the four-year-olds ($\beta = -.13$, $p = .04$).
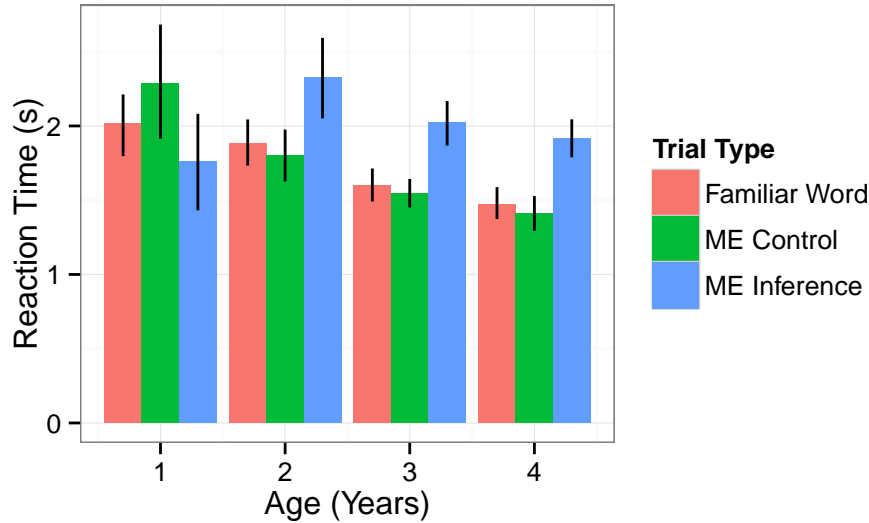
Figure 4. Reaction times in our word recognition task, plotted by age group and trial type. Error bars show 95% confidence intervals, computed by non-parametric bootstrap.

reaction times, as averages can be heavily skewed by just a small number of very long RTs.

Group-level developmental differences such as those described above do not necessarily signal a measure with high reliability. We computed split-half reliability coefficients for accuracy and reaction time based on splitting all trials using familiar words (both Familiar Word and ME Control trials) into two halves and computing the correlation between these two halves. To ensure the robustness of the statistic, we repeated this procedure with 1000 bootstrap replicates (different splits of the trials) and took the average value across these replicates. We also adjusted reliabilities using the Spearman-Brown prediction formula (?, ?) to estimate the overall reliability of the full sample of 16 items.

Reliabilities are shown in Table 3. Accuracy and reaction time show opposite patterns: while accuracy is somewhat reliable for the youngest two age groups, the presence of a ceiling effect effectively eliminates any reliability for accuracies in the older

Table 3

Mean reliability values found by a bootstrap sample of random splits of the data from familiar word trials. Values marked SB are computed by the Spearman-Brown prediction formula.

Age group	Accuracy	Reaction Time	Accuracy (SB)	Reaction Time (SB)
One-year-olds	0.45	0.09	0.61	0.05
Two-year-olds	0.31	0.34	0.42	0.44
Three-year-olds	-0.12	0.49	-0.27	0.64
Four-year-olds	0.03	0.60	0.03	0.74

two groups. In contrast, reaction times were not reliable for the one-year-olds but increased in their reliability for the older age groups. These data thus suggest that even a sample of 16 trials can yield some reliability, especially for reaction times in the older age groups. In addition, the youngest sample showed some reliability in accuracies, though we interpret this finding cautiously.

Conclusions

Mobile and tablet-based computers are quickly becoming a ubiquitous part of life—and of childhood. Children play with their parents' and their own devices, and such devices are used for education, reading, games, and video chat with friends and relatives. Many important developmental issues are raised by this ubiquity, and more research must be done to determine how, when, and what children can learn from different kinds of devices. Nevertheless, the opportunity posed by such devices for developmental psychologists is immense.

We presented a simple method for creating tablet "apps" that allows experimenters to create basic websites and use these websites to collect data on tablets. This method is substantially easier and more flexible than platform-native app development, which can be costly and time-consuming. In addition, it allows cross-platform compatibility and easy sharing for the resulting experiments. We showed in a simple word recognition experiment that this method yields reliable reaction time and accuracy data for children as young as two years of age. Data collection using tablets thus has the potential to supplement and perhaps even supplant many behavioral assessment methods currently in use.

References

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*, 505–522.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*, 48–53.

Aziz, N. A. A., Sin, N. S. M., Batmaz, F., Stone, R., & Chung, P. W. H. (2014). Selection of touch gestures for childrens applications: Repeated experiment to increase reliability. *International Journal of Advanced Computer Science and Applications*, *4*, 97–102.

Bavelier, D., Green, C. S., & Dye, M. W. (2010). Children, wired: for better and for worse. *Neuron*, *67*, 692–701.

Bertucco, M., & Sanger, T. D. (2013). Speed-accuracy testing on the apple ipad® provides a quantitative test of upper extremity motor performance in children with dystonia. *Journal of child neurology*, 0883073813494265.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*, 39–53.

Brown, A. (2011). Council on communications and media. Media use by children younger than 2 years. *Pediatrics*, *128*, 1040–1045.

Cardon, T. A. (2012). Teaching caregivers to implement video modeling imitation training via ipad for their children with autism. *Research in Autism Spectrum Disorders*, *6*, 1389–1400.

Chai, Z., Vail, C. O., & Ayres, K. M. (2014). Using an ipad application to promote early literacy development in young children with disabilities. *The Journal of Special Education*, 0022466913517554.

Christakis, D. A. (2014, May). Interactive media use at younger than the age of 2 years interactive media use at younger interactive media use at younger than the age of 2

years: Time to rethink the American Academy of Pediatrics guideline? *JAMA Pediatrics*, *168*, 399-400.

Couse, L. J., & Chen, D. W. (2010). A tablet computer for young children? Exploring its viability for early childhood education. *Journal of Research on Technology in Education*, *43*, 75–96.

Cristia, A., & Seidl, A. (in press). Touchscreen use in early childhood: Parental reports, observations, and brief intervention.

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE*, *8*, e57410.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental psychology*, *47*, 1499.

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., . . . ODoherty, K. (2010). Do babies learn from baby media? *Psychological Science*, *21*, 1570–1574.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Developmental psychology*, *37*, 630.

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *Macarthur-bates communicative development inventories: User's guide and technical manual.* Paul H. Brookes Publishing Company.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., Pethick, S., . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, *59*.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year.

Developmental psychology, *42*, 98.

Fernald, A., Pinto, J., Swingley, D., Weinbergy, A., & McRoberts, G. (1998). Rapid gains
in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*,
228.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while
listening: Using eye movements to monitor spoken language. *Developmental
psycholinguistics: On-line methods in children's language processing*, 113–132.

Frank, M. C. (2013). *ipad experiments for toddlers.* Retrieved 2013, from
`http://babieslearninglanguage.blogspot.com/2013/09/ipad-experiments-for-toddlers.html`

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential
intentions to model early cross-situational word learning. *Psychological Science*, *20*,
578–585.

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye tracking in infancy research.
Developmental neuropsychology, *35*, 1–19.

Jaeger, T. F. (2008). Categorical data analysis: Away from anovas (transformation or not)
and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy:
Effects of short-term exposure and social interaction on phonetic learning.
Proceedings of the National Academy of Sciences, *100*, 9096–9101.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary
knowledge in infancy predict cognitive and language outcomes in later childhood.
Developmental science, *11*, F9–F16.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to
constrain the meanings of words. *Cognitive Psychology*, *20*, 121–157.

Markman, E. M., Wasow, J., & Hansen, M. (2003). Use of the mutual exclusivity

assumption by young word learners. *Cognitive Psychology*, *47*, 241–275.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making*, *5*, 411–419.

Radich, J. (2013). Technology and interactive media as tools in early childhood programs serving children from birth through age 8. *Every Child*, *19*, 18.

Rideout, V. (2014). Learning at home: Families' educational media use in America. *New York: The Joan Ganz Cooney Center at Sesame Workshop*.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! socially contingent interactions help toddlers learn language. *Child development*, *85*, 956–970.

Rvachew, S. (2014, March). *Are baby videos bad for babies?* Retrieved from `http://digitalmediaprojectforchildren.wordpress.com/2013/09/22/are-baby-videos-bad-for-`

Shuler, C., Levine, Z., & Ree, J. (2012). ilearn ii: An analysis of the education category of apples app store. In *New york: The joan ganz cooney center at sesame workshop.*

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American scientist*, 421–457.

Waddington, H., Sigafoos, J., Lancioni, G. E., OReilly, M. F., van der Meer, L., Carnett, A., ... others (2014). Three children with autism spectrum disorder learn to perform a three-step communication sequence using an iPad-based speech-generating device. *International Journal of Developmental Neuroscience*.

Zack, E., Barr, R., Gerhardstein, P., Dickerson, K., & Meltzoff, A. N. (2009). Infant imitation from television using novel touch screen technology. *British Journal of Developmental Psychology*, *27*, 13–26.