

# Young Children's Performance on Self-administered iPad Language Activities

Jared Bernstein<sup>1</sup>, Ognjen Todic<sup>2</sup>, Kayla Neumeier<sup>3</sup>, Katharyn Schultz<sup>3</sup>, Liang Zhao<sup>3</sup>

<sup>1</sup>Tasso Partners, Palo Alto, California, USA, <sup>2</sup>Keen Research LLC, Mill Valley, California, USA

<sup>3</sup>Knowledge Technologies, Pearson, Menlo Park, California, USA

jared.bernstein<>Stanford.edu, Schultz.katharyn<>gmail.com

## Abstract

Twenty-nine English language activities were implemented on a touch-tablet computer. Some activities were focused on a single skill (e.g. reading or speaking), while others involved several integrated skills (e.g. listening and writing). Materials were presented in several modalities, including speech only, speech with figure, silent video, text, speech with text, and speech with text and figure. Response modalities included speech, typing, touch, dragging screen objects, selecting and/or arranging words, and drawing figures. Various test-like sequences of 24 to 45 items were presented to 784 children, 53% from non-English speaking homes. Analysis of over 28,000 responses to these self-administered activities indicates that most activities can be successfully modeled by a single short instructional video example. By age 8 years, nearly all children will respond meaningfully to about 95% of these specific activities. Examples of these activities and child responses are presented.

**Index Terms:** Language, assessment, touch tablet, four skills, ELL, children, modeling, modality.

## 1. Introduction

The design and development of instructional systems often proceeds by trial and error. For example, in the United States, various committees formulate standards that describe activities students should be able to perform (e.g. “*Students at all levels of English language proficiency will evaluate [an] author’s bias.*”), giving finer definitions for different levels, ages, or grades (<http://www.wida.us/standards/eld.aspx>). Publishers and system developers then design instructional material and tests that align with these standards, even when standards have not been developed with a view to making the defined skills easily measurable. This situation can leave test developers searching for methods that will elicit scoreable samples of student performance within feasible bounds of time and cost. If we understood which kinds of computer-based activities are intuitive for children, we could more accurately plan the development of computer-based assessments and instructional systems.

At the same time, a remarkable proliferation of smart phones and touch-tablet computers, like the Apple iPhone and iPad, and similar devices, has prompted interest in using these devices in education. Touch-tablet computers have touch-sensitive screens, virtual keyboards, accelerometers, and good quality audio I/O. Children are excited to use touch tablets, and their availability opens new possibilities in elicitation and capture of student performance with first and second language, as well as other areas including mathematics. We report an experiment to identify the touch-tablet-enabled presentation and response modalities that children can handle successfully (in 2012 in North America) without any adult guidance or instruction.

## 2. Approach

Streeter et al. (2011) organize most current applications of technology in language assessment within the grid of Table 1, however touch tablets and smart phones allow several extensions not foreseen in that table of possibilities.

Table 1. *Traditional presentation and response modalities.*

| Scoring Focus         | Presentation |         | Response |         |
|-----------------------|--------------|---------|----------|---------|
|                       | Spoken       | Written | Spoken   | Written |
| Declarative Knowledge | +            | +       | +        | +       |
| Language Skills       | +            | +       | +        | +       |

In the present study, a set of 29 different language activities (or test item types) were implemented on an iPad-2 touch-tablet computer. These items were designed to cover many combinations of input and output modalities that are available on a touch-tablet, such that the performance of a young child (age 4-11 years) can, in principle, be measured automatically from the responses. Some activities were designed to elicit information about a single skill (e.g. just reading or speaking), while others elicited performances that reflected several integrated skills (e.g. both listening and writing). Materials were presented in several modalities, including speech only, speech with figure, silent video, text, speech with text, and speech with text and figure. Response modalities included speech, typing, touch, dragging screen objects, selecting and/or arranging words, and drawing figures. Within each activity type, among 15-20 items, certain items were designed for either first or fifth graders (aged 5-6 or 9-10), while other items were designed for use with both groups. Our first analysis of the children’s responses was designed to answer three questions:

- (1) Which activities can children understand well enough (without adult help) to perform meaningfully on?
- (2) Which specific activities yield the most information about a child’s relative language skills, and which materials are most appropriate for ages 4-7 and 8-11?
- (3) Which activities best discriminate English language learners (ELLs) from ‘mainstream’ students?

### 2.1. Materials

Apple’s iPad was used as the experimental delivery platform because it had a rich Software Development Kit that easily supports multimedia presentation and accepts a number of different gesture controls. A test-like presentation flow was implemented on the iPad, as shown in Figure 1.

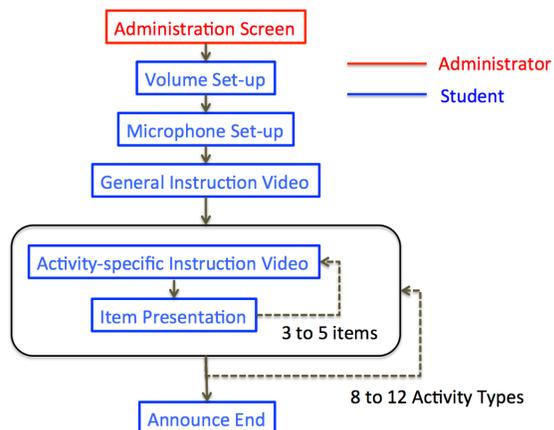


Figure 1: Presentation-flow of items within Activity Types.

Twenty-five of the Activity Types studied are listed in Table 1. Some isolate a receptive skill (listening or reading) by using a non-linguistic response mode. The Silent RT Description isolates speaking by presenting a silent video clip (8-15 seconds) for the student to narrate or describe in real time (RT).

Table 2. Activities, with modalities and language skills (Listening, Speaking, Reading, Writing, Usage).

| Activity Types (Information Rank) | Elicit        | Respond | Skills |
|-----------------------------------|---------------|---------|--------|
| Repeat                            | Voice         | Speak   | L,S    |
| Silent RT Description (3)         | Silent Video  | Speak   | S      |
| Oral Read & Answer (1)            | Text          | Speak   | R,S,L  |
| Situated Polite Request (4)       | Voice, Figure | Speak   | L,S,U  |
| Seeded Sentence                   | Voice, Text   | Speak   | R,S    |
| Passage retell                    | Text, Voice   | Speak   | R,S    |
| Teacher Teach                     | Video, Voice  | Speak   | L,S    |
| Hear-Touch                        | Voice, Figure | Touch   | L      |
| Hear-Point Conversation           | Voice, Figure | Touch   | L      |
| Hear-Move                         | Voice, Figure | Gesture | L      |
| Hear-Path                         | Voice, Figure | Gesture | L      |
| Anaphora                          | Text          | Touch   | L,R    |
| Read-Move                         | Text          | Gesture | R      |
| Place Words                       | Text          | Gesture | R      |
| Passage Read & Answer             | Text          | Type    | R,W    |
| Non-word Read Aloud               | Text          | Speak   | R,S    |
| Recognize Word                    | Text          | Touch   | R      |
| Recognize Non-word                | Text          | Touch   | R      |
| Insert Adverbial                  | Text          | Gesture | R,W,U  |
| Word Choice                       | Text          | Gesture | R,W,U  |
| Sentence Cloze                    | Text          | Type    | R,W,U  |
| Find Error                        | Text          | Touch   | R,U    |
| Spell Words (5)                   | Voice, Figure | Type    | W      |
| Gr1 Passage, write-word           | Text          | Type    | R,W    |
| Gr5 Passage re-write (2)          | Text          | Type    | R,W    |

For example, in a **Hear-Point Conversation**, the student hears a conversation between two voices, during which two or more concrete objects in the figure are mentioned in passing. The student has been instructed to touch any object as it is mentioned in the running conversation. See Figure 2.

## 2.2. Procedures

Each child sat before an iPad on a table-top stand, in a small room. Typically 3 to 8 children were doing the activities individually in the same room, but not in synchrony with each other. The audio was presented via the iPad’s built-in speaker and spoken responses were recorded through the built-in microphone. After an administrator entered the child’s ID and selected an age-appropriate version to run, students were left alone to figure out what the task demands were. They had been told that the iPad gave a test, but most referred to it as a game. Each student encountered a sequence of activities, with 3 to 5 items of each activity type presented, following a single 10-20 second video showing a child doing a sample item (see Figure 1). Most children aged 4-7 encountered 36 items, while those aged 8-11 most often had 43 items to do.



Figure 2: Example Hear-Point Conversation figure. Several objects are touchable (e.g. dad, chair, cake, boy, fork), although only two are mentioned in the audio conversation.

## 3. Results

In July-August 2012, we ran 326 students; then we revised some items, introduced several new activities, and in October 2012, ran another 458 students (N=784). Almost all students who participated were from low-income homes, with 40% currently in official ELL status, but only 47% listed English as their home language. Specific results:

- (1) Which activities work? 27 of 29 activities elicited useful responses from > 85% of students aged 8-11. Among students 4-7, 24 of 29 activities elicited useful responses from at least 75% of the students.
- (2) Which yield the most skill information? The best are Oral Read Passage, Situated Polite Request, and Silent Video RT Description, with traditional Non-word Read Aloud and Spell Word nearly as good.
- (3) Which discriminate ELLs from other students? Items types yielding the most information are noted in Table 2. Across all activities and ages, the ELL-nonELL score difference averaged about 3% of the observed score range. The items that best discriminated ELLs were Repeats, Read-Moves, Hear-Moves, Polite Requests, and Hear-Paths.

In conclusion, most activities can be successfully modeled by a single short video example. In this U.S. sample, by age 8 years, children respond meaningfully to almost all these items about **95%** of the time, regardless of first language.

## 4. References

[1] L.Streeter, J.Bernstein, P.Foltz, & D.DeLand, D. (2011). *Pearson’s Automated Scoring of Writing, Speaking and Mathematics*. (pdf) at <http://kt.pearsonassessments.com>