Validation of an Expressive and Receptive Tablet Assessment of Early Literacy

Michelle M. Neumann^{1, 2}, Sheena Worrall ³, & David L. Neumann³

¹Griffith Institute for Educational Research, Griffith University, QLD, Australia

²School of Education and Professional Studies, Griffith University, QLD, Australia

³School of Applied Psychology, Griffith University, QLD, Australia

Author Notes

Please direct correspondence to: Michelle Neumann: School of Education and Professional Studies, Griffith University, Gold Coast QLD 4222, Australia. Telephone: +61(0)7 55529785
E-mail: m.neumann@griffith.edu.au

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Abstract

Touch screen tablets are used in the classroom for assessment. Little is known about the psychometric properties of tablet based assessments. This study examined the validity and reliability of an expressive and receptive assessment app designed to measure literacy skills. Children ($N = 45$; 3-5 years) completed the app assessments for alphabet and word skills using a tablet. Children ($N = 15$) were post-tested 1 to 2 weeks later to check test score stability. The tablet measures showed good internal consistency (all $\alpha > .90$) and test-retest reliability (ICC range = .87 to .98). Agreement between scores for the tablet and paper-based tests was also high (ICC range = .91 to .98). This indicates that expressive and receptive tablet tests provide valid and reliable measures.

Key words: Touch screen tablets, assessments, expressive test, receptive test, literacy, children

Introduction

Touch screen tablets and apps are being increasingly used in classrooms for learning and teaching (Marsh et al., 2015; Ralph & Petrina, 2019; Woloshyn, Bajovic, & Worden, 2017), particularly in curriculum areas such as literacy (Neumann, 2018; Oakley, Wildy & Berman, 2018; Xie, Peng, Qin, Huang, Tian, & Zhou, 2018; Herodotou, 2018; Neumann & Neumann, 2014a). The touch-based operational features of tablets do not rely on mouse and keyboard skills and allow young children easy operation of these devices using their fingers (Crescenzi, Jewitt, & Price, 2014; Fletcher-Watson, 2013; Goodwin & Highfield, 2012). Further, their interactive and multimodal features provide opportunities for communication, ownership of learning, social interaction, play, and motivation (Flewitt, Messer, & Kucirkova, 2015; Geist, 2014; Lee, 2015). They also have the potential to motivate, engage and support children in assessment tasks (Woloshyn et al., 2017). However, little is known about the validity and effectiveness of using tablets and apps to assess learning in early years classrooms (Bennett, 2011; Carson, Boustead, & Gillon, 2015; Carson, 2017; Lu, Y., Ottenbreit-Leftwich, Ding, Ai-Chu., & Glazewski, 2017; Mohamad, Lakulu, & Samsudin, 2016; Neumann & Neumann, 2018; Woloshyn et al., 2017).

A key theory underpinning assessment argues that it is a social practice or activity (Broadfoot & Black, 2004; Looney, Cumming, Van der Kleij, & Harris, 2018; Silseth & Gilje, 2019). From this sociocultural perspective, assessment is an interactive, dynamic and continuous process where students actively construct knowledge (Heritage, 2018). It is embedded in the teaching and learning process with this process mediated by cultural tools, driven by advances in technology. Furthermore, assessment processes are closely linked with the social interaction of learners and the construction of knowledge achieved by a novel-expert relationship where scaffolding is provided to extend students to achieve their learning

EXPRESSIVE AND RECEPTIVE TABLET TESTS

potential. As such, assessment tools that are underpinned by sociocultural perspectives are likely to benefit student learning (Hodges, Eames, & Coll, 2014).

Assessment apps used via tablets have the potential to be powerful assessment tools for teachers and students (Neumann & Neumann, 2018; Woloshyn et al., 2017). The use of multimodal assessment apps (Carson, Gillon, & Boustead, 2011; Neumann & Neumann, 2018) can provide flexibility of task presentation (e.g., *receptive response mode* such as tap the screen or *expressive response mode* such as speak into an inbuilt microphone) to suit teacher's assessment needs or children's capabilities. Assessment apps can be designed to efficiently track learning progress and automated data scoring protocols can guide teachers in lesson planning (Lu et al., 2017). This increased efficiency can free up more time for teachers to work directly with students and tailor their teaching to individual student needs (Koomen & Zoanetti, 2016).

Due to increasing academic expectations, numerous tests have been designed and standardized for classroom use such as the Phonological Awareness and Literacy Screenings – PreK (PALS and Test of Early Reading Ability, 3rd Edition (TERA-3). However, as reported in a review of 26 assessment tools by Carson (2017), few have been designed for use on a tablet and app platform. Of those assessment apps, no app was found to administer the test items itself. Instead, the teacher was required to ask the test questions (presented on the screen) and manually score and collect the test data. These findings demonstrate the need to design interactive apps that allow young children to complete assessment tasks independently. Moreover, it is important to test the validity and reliability of new assessment apps and to test their effectiveness and efficiency to collect, automatically score, summarise, interpret, and report on data (Carson et al., 2011; Carson et al., 2015).

The efficiency of using a computer-based assessment has been demonstrated by Carson et al. (2011) who administered a laptop version of a letter knowledge test to 33

EXPRESSIVE AND RECEPTIVE TABLET TESTS

children aged 4 to 5 years. Letters were present in a 3×2 grid on the screen and the children were asked questions like “show me the letter m” or “show me the letter that makes the /m/ sound”. The children pointed to a letter on the screen and the teacher used a mouse to click on the letter indicated by the child. A comparable paper-pencil administered test was also conducted and it was found that the computer-based assessment was 20% faster to administer than the paper-pencil test version. This highlights the potential advantages of using digital assessment tools in the classroom. Furthermore, Carson et al. (2015) tested a computer-based phonological awareness test in 95 children aged 5 years. Although little detail on administration of the test was provided, the findings overall showed that the computer-based test had good content validity and allowed the differentiation of children’s literacy skills on entry to school.

The studies conducted by Carson et al. (2011, 2015) not only highlight the potential benefits of computerized assessments, but also the importance of design features. The children did not themselves directly interact with the assessment app. It was the teacher who operated the mouse to click on the answer. As noted, a touch screen tablet is easier for a young child to operate. For a child to make a response on a tablet themselves, they can complete tests that are designed to be either receptive (students listen to a narrated instruction and respond by tapping an answer on the screen) or expressive (students orally respond to the verbal question given by the app) (Neumann & Neumann, 2018). Receptive language skill can be defined as the ability to understand letters or words when they are read or heard, while expressive language skill is being able to mentally retrieve words and say them out loud (Vatalaro, Culp, Hahs-Vaughn, & Barnes, 2018). Assessments that can provide both separate receptive and expressive tests may provide a more accurate assessment of the extent of student knowledge and understandings (Wise, Sevcik, Morris, Lovett, & Wolf, 2007).

EXPRESSIVE AND RECEPTIVE TABLET TESTS

A recent study by Neumann and Neumann (2018) tested the validity and reliability of a receptive test for preschool children's ($N = 99$, M age = 4.83 years) alphabet (upper and lower-case letter name and sound), word reading, and numeral skills administered using a touch screen tablet. The assessments used a receptive test format in which four items were presented on a tablet screen and a narrator voice asked children to touch a specific target item (e.g., "touch the letter P"). Scores on all the receptive test measures showed excellent internal consistency (all $\alpha > .94$), good test-retest reliability ($r_s > .84$), and good agreement with the same literacy tests administered using traditional paper-based methods ($r_s > .79$). The findings demonstrated that the tablet-based assessment approach had sound psychometric properties. In addition to the psychometric evaluation, the study revealed that the assessment app has the potential to make data collection more efficient than the traditional paper-based assessment. The design features of the app such as age appropriate operation (e.g., tap) and motivational features (e.g., animated smiley face) engaged children in the task and they enjoyed completing the assessments.

However, the assessment approach used by Neumann and Neumann (2018) was limited because the assessment app only provided a receptive test of literacy skills. The app has since been revised to include an expressive test of literacy skills in addition to the receptive test. To assess children's expressive literacy knowledge on a tablet, speech recognition software is required. Raskind and Higgins (1999) showed that computer-based speech recognition technology could be used with students with learning disabilities to significantly improve their word recognition and reading. Baker (2016) found that iPad speech recognition technology had progressed to a point where it could be used to help struggling first-grade readers by transcribing their dictated stories into text. However, Baker (2016) also reported that the story writing software used in the classroom sometimes made errors in transcribing children's oral language, including the letter *a* consistently being

EXPRESSIVE AND RECEPTIVE TABLET TESTS

transcribed as *hey*, and *hill* being transcribed as *hell*. In spite of this, Baker's (2016) study found that it was feasible to support young readers' literacy development with a speech-recognition-enabled tablet app. However, few of these types of voice recognition and expressive studies with tablets and apps have been conducted and to our knowledge no study has examined tablet-based assessment apps that uses speech-recognition technology and an expressive tablet assessment with young children.

The Present Study

In order to further explore the benefits of using of tablet-based assessment tools to measure children's learning it is important to test the validity and reliability of new digital assessments for classroom use (Carson et al., 2015; Carson, 2017; Neumann & Neumann, 2018; Woloshyn et al., 2017). The assessment app designed for the present study contained similar features as the assessment app used in Neumann and Neumann's (2018) study. The app design features were adapted from app design criteria recommended by Hillman and Marshall (2009), Miller and Kocurek, (2017), and Marsh (2016). The present assessment app contained the following design features: developmentally appropriate (targeted to young children and is challenging but not frustrating with age appropriate operational features e.g., tap); underpinned by social-cultural theory where children play an active role in their learning and assessment experiences; provides autonomous decision making opportunities; feedback is easy to interpret and positive to guide the child's performance rather than as a win or lose outcome; interactive features allow active child participation and has flexible response mode features (e.g., contains the use of external speech recognition software that can help decipher and collect individualized responses). For the test administrator, the assessment app was designed so that the app automatically scored and recorded the data which could be later accessed through a cloud-based database for interpretation and reporting.

EXPRESSIVE AND RECEPTIVE TABLET TESTS

The present study specifically focussed on examining two design features of the assessment app (receptive and expressive test response). It compared the validity and reliability of an online expressive and receptive tablet-based assessment tool to extend previous research (Carson et al., 2011; Carson et al., 2015; Neumann and Neumann, 2018). A comparison of the tablet tests against traditional paper-based tests was also conducted to provide a more rigorous analysis of each of these test formats. The measures selected for this study were early literacy skills (alphabet and word knowledge) in preschoolers because limited research exists on early literacy tablet assessments and children (Carson et al., 2011; Carson et al., 2015). It was hypothesised that the findings of Neumann and Neumann (2018) would be replicated and that good reliability and validity would be found for the receptive tablet measure, as shown by good test-retest reliability, internal consistency, and agreement with a paper-based expressive assessment. In addition, it was hypothesised that the expressive test format would also demonstrate acceptable reliability and validity. This would also be shown by strong test-retest reliability, internal consistency, and agreement with a paper-based expressive test.

Method

Participants

Twelve preschools in South-East Queensland and Northern New South Wales, Australia agreed to participate in this study. Parents of 51 English-speaking children (32 males and 19 females) completed consent forms, allowing their child to participate in the study. Verbal consent was also obtained from each of the children themselves, immediately prior to conducting the assessments. Participating children were aged from 3 to 5 years ($M = 4.65$ years; $SD = 0.47$; range = 3.65-5.42 years). One child was excluded from the study due to unforeseen environmental noise in the preschool that prevented the child from hearing the instructions clearly and five children were excluded due to developmental delays or speech

EXPRESSIVE AND RECEPTIVE TABLET TESTS

and hearing disorders, as reported by parents, leaving a final sample of 45 children (29 boys and 16 girls). Based upon parental education level and occupation, socioeconomic status (SES) was calculated using the Hollingshead (1975) 4-factor index (range 8-66). The mean SES was in the middle SES range ($M = 36.64$; $SD = 12.13$; range 11-61). Most families (96.1%) reported having one or more tablet devices at home ($M = 1.82$; $SD = 1.10$; range 0-5), with 76.5% of families reporting that their child used a tablet device daily ($M = 48.86$; $SD = 55.66$; range 0-240 minutes).

Procedure

Approval to conduct the study was obtained from the University Ethics Committee. Parents of participating children completed the consent form and parent questionnaire at home. The child assessment sessions were conducted in a room at the child's preschool. A random selection of 15 children were retested ($M = 11.6$ days following the initial tests) to examine temporal stability of the measures. An assessment account and unique identifier code for each participating child was created in the app. The app was run on an iPad and contained three multiple choice receptive tests (upper case letter name, lower case letter name, word reading) and three expressive tests (upper case letter name, lower case letter name, word reading). The assessment also included a paper-based expressive test (upper case letter name, lower case letter name, word reading). Administration time for the whole assessment battery was approximately 30 minutes. The order of individual alphabet letters and words within the tablet tests were automatically randomized by the app's software. To account for any potential learning effects amongst the participants, a counter-balanced design was used for the presentation order of the receptive tablet, expressive tablet, and expressive traditional assessment methods.

Measures

Literacy assessment app. The tablet literacy assessment app used for both receptive and expressive literacy tests was an updated version of the receptive app developed by Neumann and Neumann (2018) and Neumann (2016). The app was downloaded onto an iPad Pro [512GB; 305.7 mm (h) x 220.6 mm (w) x 6.9 mm (d)] device. The app includes a set of receptive and expressive tests for uppercase letter name (out of 26), lowercase letter name (out of 26), and word knowledge (out of 10; the, look, and, come, will, he, up, here, to, me). The app was used for both the initial test assessments and the post-test assessments. Each test began with a pop-up orientation task page where a picture of a frog appeared and was followed by the app's narrator voice asking the child to "Touch the frog". This interactive task served to confirm that the child can follow the instruction to tap on an item using their finger.

For the receptive tablet assessment tasks, a quadrant array of four items is presented. Screenshots are shown in Figure 1 for examples of uppercase letters and sight words. Each target letter or word is presented in a randomised position in the array. Three other non-test options (distractors) were presented randomly in the other three quadrants. The narrator's voice provides the instruction (e.g. "Touch the letter Q"). The child responds by using their finger to tap the letter or word. A short tone is heard when a letter or word is touched. No feedback is provided to indicate a correct or incorrect response.

Figure 1. Examples of the receptive test interface for uppercase letter name (left panel) and word reading (right panel).

For the tablet literacy expressive tests, the app uses a voice-recognition response format. As shown in Figure 2, the target letter or word is presented accompanied by the app narrator's instruction to "say this letter" or "say this word". Tapping the circular microphone icon activated the speech recognition feature of the app. Prior to starting the expressive test, the child was instructed by the researcher to (a) identify the letter or word, (b) touch the microphone icon, and (c) and say their answer. If the child did not know the letter or word, the researcher gave the instruction to touch the microphone button, to allow the app to proceed to the next item on the subtest. During each of the expressive tests (upper case letter name, lower case letter name, and word reading), a separate audio recording of each child's

EXPRESSIVE AND RECEPTIVE TABLET TESTS

response was also made using a Olympus Digital Voice Recorder WS-833, to cross-validate the accuracy of the speech-recognition technology.

Figure 2. Examples of the expressive tests interface, for uppercase letter knowledge (left panel) and word reading (right panel).

Results for the receptive and expressive tests were stored on a secure online database and accessed by using a web browser. The results were presented on a dashboard that contained the date, time, and test score. A count of correct responses for each of the three test types (upper case letter name, lower case letter name, and word reading) was used for comparisons between the expressive tablet tests, receptive tablet tests, and the paper-based test. However, the receptive tablet test was multiple choice and required children to select the correct answer from the four items presented. If a child was unsure about the correct response, they could select an item based on their best guess. Chance guessing has the

EXPRESSIVE AND RECEPTIVE TABLET TESTS

potential to inflate scores when a simple count is used for multiple option tests. A retrospective correction for guessing was applied to counts of correct scores by using the formula reported by Frary (1988) and recommended by Lesage, Valcke and Sabbe (2013): $FS = R - W / (C - 1)$. This formula score (FS) is a count of right responses (R) after a deduction has been made which takes into account the number of wrong responses (W) and the number of choices (C) for each question (for the tablet receptive test $C = 4$). The FS will produce a score of zero if a student's guesses are completely random, and if more incorrect answers are made than expected by chance, the FS score may register a negative score. As negative values are not possible with paper-based literacy assessments, any negative scores are changed to zero.

The tablet assessment app's expressive letter and word knowledge response format required the child to say their response while the app was in record mode (i.e., when the microphone icon was touched). If a child was unsure of the answer, they had the option of either guessing, saying "I don't know" or something similar, or not responding. The researcher listened back to the digital recordings of responses and compared them to the app's scores for that test to confirm a count of correct responses for each expressive test for that participant. A simple count of correct responses was chosen as the most appropriate measure for this set of tests because there were no multiple choice response options given in the expressive test.

Paper-based literacy assessments.

Upper case letter name, lower case letter name, and word reading skills were also measured using an expressive paper-based assessment, as used by previous studies (Drouin, Horner, & Sondergeld, 2012; Neumann & Neumann, 2014b). This method has also been used as a comparison test for studies examining technology-based early literacy assessment methods and uses an expressive test format (Carson et al., 2011; Neumann and Neumann, 2018; Piasta, Petscher, & Justice, 2012).

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Letter name knowledge. Each child was shown white cards (8 × 8 cm) each presented in a random order with an upper or lower case letter printed in Century Gothic size 120 font. As each upper case letter was presented, the researcher asked “What is the name of this letter?” The test was scored 1 for a correct response and 0 for an incorrect response (Max = 26). With a pencil, the researcher recorded each response on a score sheet that listed all the 26 upper case letters. This process was repeated for lower case letter name knowledge (Max = 26).

Word reading. Children were shown a list of 10 words (the, look, and, come, will, he, up, here, to, me) printed on a sheet of white paper. The researcher pointed to each word in turn and asked “Can you read this word?” Children were scored 1 for a correct response and 0 for an incorrect response (Max score = 10). With a pencil, the researcher recorded each response on a score sheet containing the list of words.

Questionnaire. The parent questionnaire was used to collect demographic information (child age, parent education level and occupation, digital device ownership and the time the child spent using tablets each day) and details of child development problems.

Results

The descriptive statistics are presented in Table 1 and 2. Preliminary analyses examined the distributions of scores for normality and outliers, for the expressive, receptive, and paper-based measures. The normality of the data distributions and any outliers were determined by using histograms, boxplots, and skew and kurtosis scores ($p < .05$). Standardized measures of skew, kurtosis and the Shapiro Wilke test of normality for the 12 variables showed only three tests lacked normality. Word reading paper-based expressive, tablet expressive, and tablet receptive formula score tests showed significant positive skew due to a zero-inflated floor effect. Although children could produce an utterance or tap the screen in response to the word items, a high proportion of these utterances were incorrect. For

EXPRESSIVE AND RECEPTIVE TABLET TESTS

the total count of correct responses for paper-based word reading expressive test, 86.67% of the test items were incorrect. For the tablet word reading expressive subtest, 88.89% of the test items were incorrect. For the guessing-adjusted tablet word reading receptive formula score subtest, 53.33% of responses were incorrect. These findings reflect the young age of the sample ($M = 4.65$ years) and limited word reading ability. Given the high number of zero scores for the word reading tests, the frequencies for these tests are reported in Table 1 and no further analyses were conducted for this measure.

Table 1

Frequency of Scores for Paper-based and Tablet Tests for Word Reading ($N = 45$)

Measure	Frequency											
	0	1	2	3	4	5	6	7	8	9	10	
Paper-based Expressive	39	3	0	0	0	1	1	0	0	0	1	
Tablet Expressive	40	2	0	0	1	0	1	0	0	0	1	
Tablet Receptive	7	5	9	10	6	1	3	1	1	0	2	

Although not significant, there was a general trend of negative kurtosis across all the measures, except for the word reading receptive count correct test, which had significant negative kurtosis. Descriptive statistics calculated included mean scores, range, and standard deviations (see Table 2). However, because of concerns that the negative kurtosis might affect the mean and standard deviations, the median and interquartile ranges were also presented.

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Table 2

Descriptive Statistics for the Paper-based and Tablet Tests for each

Measure (N = 45)

Literacy Measure	Min	Max	Mean	SD	Median	IQR
<i>Upper Case Letter Name</i>						
Paper-based expressive test	0.00	26.00	13.71	8.81	14.00	16.00
Tablet expressive test count correct answers	0.00	26.00	13.58	8.92	14.00	18.50
Tablet receptive test count correct answers	5.00	26.00	19.51	6.39	22.00	10.50
Tablet receptive test formula score	0.00	26.00	17.41	8.39	20.67	14.00
<i>Lower Case Letter Name</i>						
Paper-based expressive test	0.00	26.00	12.20	8.18	12.00	14.50
Tablet expressive test count correct answers	0.00	26.00	11.87	8.41	12.00	14.75
Tablet receptive test count correct answers	4.00	26.00	17.64	6.61	19.50	11.25
Tablet receptive test formula score	0.00	26.00	15.06	8.39	17.33	15.00
<i>Word Reading</i>						
Paper-based expressive test	0.00	10.00	0.53	1.85	0.00	0.00
Tablet expressive test count correct score	0.00	10.00	0.49	1.80	0.00	0.00
Tablet receptive test count correct score	0.00	10.00	3.00	2.47	3.00	3.00
Tablet receptive test formula score	0.00	10.00	1.53	2.56	0.67	2.00

Validity and reliability

Internal consistency for each literacy test was tested with Cronbach's α . Analyses indicated that internal consistency was excellent for all paper-based and tablet expressive and receptive upper and lower case letter name tests. Moreover, the values for all tablet tests were comparable to those obtained using the paper-based tests. Results were as follows: paper-based upper case letter name expressive test, $\alpha = .96$; tablet upper case letter name expressive test, $\alpha = .96$; tablet upper case letter name receptive test, $\alpha = .92$; paper-based lower case letter name expressive test, $\alpha = .95$, tablet lower case letter name expressive test, $\alpha = .95$; and tablet lower case letter name receptive test, $\alpha = .91$.

The level of agreement between the paper-based expressive tests and the tablet tests was calculated using the intraclass correlation coefficient (ICC), to assess their convergent validity. This coefficient is recommended by Watson and Petrie (2010) for assessing the reproducibility of paired numerical datasets. A strength of this approach is that the ICC can estimate reliability for a single assessor and can differentiate between relative consistency and absolute agreement (Streiner & Kottner, 2014). When using the ICC to determine the degree of equivalence between a score on one scale and a score on another scale plus systematic error, Li and Koo (2016) recommend a 2-way mixed-effects model. All data in the current study was collected by the same researcher using the same paper-based methods, iPad and sets of tablet tests for each participant, so the 2-way mixed-effects model analysis with consistency type was adopted for the convergent validity check. Interpretation of ICC values was based on the 95% confidence interval.

As shown in Table 3, an excellent degree of reliability was found between paper-based and tablet expressive measures for upper case and lower case letter name. Consistency for the comparison between the receptive upper case letter name count of correct answers and

EXPRESSIVE AND RECEPTIVE TABLET TESTS

the tablet formula score was excellent, as was the agreement between the count of correct answers for lower case letter name and the tablet formula score for lower case letter name.

Table 3

Intraclass Correlation Coefficients (95% CI) for the Agreement between Paper-based and Tablet Tests for each Measure (N = 45)

Measure	Tablet count of correct answers	Tablet formula score
<i>Expressive</i>		
Upper case letter name	.98 [.97-.99]	
Lower case letter name	.98 [.97-.99]	
<i>Receptive</i>		
Upper case letter name	.91 [.83-.95]	.93 [.88-.96]
Lower case letter name	.92 [.85-.95]	.93 [.87-.96]

Note: The tablet formula score is only relevant for receptive tests.

The means for the tests for upper case and lower case letter name were also examined for agreement. This analysis examined the paper-based expressive tests, the tablet expressive and receptive tests, and the tablet formula score receptive tests. As shown in Table 2, the count of correct responses for the receptive tests produced higher mean values than the receptive test scores that were corrected for guessing using the formula score. A one-way repeated measures ANOVA tested whether the differences between the tests were statistically significant. Due to significant violations of the sphericity assumption, Greenhouse-Geisser corrections were used (Field, 2009). For the comparison of the letter name tests, the ANOVA was statistically significant for both upper case letter name tests, $F(1.55, 68.31) = 51.09, p < .001, \eta_p^2 = .54$, and lower case letter tests, $F(1.52, 65.13) = 44.09, p < .001, \eta_p^2 = .51$.

All test scores were compared amongst each other, using pairwise *t*-tests, with Bonferroni-adjusted degrees of freedom ($\alpha' = .0083$). No significant difference was found between the upper case letter name paper-based expressive and tablet expressive subtests, *t*

EXPRESSIVE AND RECEPTIVE TABLET TESTS

(44) = 0.38, $p = .704$, $d = 0.02$. Significant differences were found between the following five upper case letter name measures: paper-based expressive and tablet receptive, $t(44) = 8.62$, $p < .001$, $d = 0.75$; paper-based expressive and tablet receptive (formula score), $t(44) = 5.72$, $p < .001$, $d = 0.43$; tablet expressive and tablet receptive, $t(44) = -8.62$, $p < .001$, $d = 0.76$; and tablet expressive and tablet receptive (formula score), $t(44) = 5.82$, $p < .001$, $d = 0.44$. In each case, the receptive tests showed a higher score than the expressive tests. Furthermore, the upper case letter name tablet receptive test score was significantly higher than the upper case letter name tablet receptive formula subtest score, $t(44) = 7.00$, $p < .001$, $d = 0.28$.

For the comparisons among the lower case letter name tests ($\alpha' = .0083$), no significant difference was found between the paper-based expressive and tablet expressive tests, $t(44) = 1.05$, $p = .298$, $d = 0.04$. Significant differences, again indicating higher scores for receptive tests, were found for the lower case letter name tests for paper-based expressive and tablet receptive, $t(43) = 8.41$, $p < .001$, $d = 0.70$; paper-based expressive and tablet receptive formula score, $t(43) = 4.16$, $p < .001$, $d = 0.32$; tablet expressive and tablet receptive, $t(43) = 8.48$, $p < .001$, $d = 0.74$; and tablet-based expressive and tablet receptive formula score, $t(43) = 4.56$, $p < .001$, $d = 0.35$. The lower case letter name tablet receptive and tablet receptive subtest score was higher than the formula subtest scores, $t(43) = 9.13$, $p < .001$, $d = 0.34$.

Test-retest reliability

The ICC was also used to test stability for all measures over time (i.e. test-retest reliability). As shown in Table 4, test-retest reliability was good for all measures (ICCs $\geq .87$). Upper and lower case letter name paper-based and tablet-based expressive tests showed excellent reliability between Time 1 and Time 2. While reliability coefficients for the tablet

EXPRESSIVE AND RECEPTIVE TABLET TESTS

count of correct answers, receptive tests for upper case and lower case letter knowledge were lower than those of the expressive tests, they were still in the very good range.

Table 4

Intraclass Correlation Coefficients (95% CI) for the Consistency between Time 1 and Time 2 for the Paper-based Tests and Tablet Tests for each Measure (N = 15)

Measure	Paper-based expressive test	Tablet count of correct answers expressive test	Tablet count of correct answers receptive test	Tablet formula score receptive test
Upper case letter name	.98 [.93-.99]	.99 [.95-1.00]	.88 [.64-.96]	.88 [.64-.96]
Lower case letter name	.98 [.95-.99]	.97 [.92-.99]	.87 [.63-.96]	.87 [.63-.96]

Relationships among the tablet tests and with demographic variables

The relationships between upper and lower case letter name expressive and receptive knowledge and the demographic variables of age and SES were examined. Spearman's rho was used, due to the ordinal scale of SES and the correlations which can be seen in Table 5. SES also showed a strong positive correlation with upper and lower case expressive tablet test and the lower case receptive formula score test scores (all $r_s > .39$, $p_s < .01$). The upper case receptive formula score was also positively correlated with SES ($r_s = .36$, $p < .05$). Age showed a positive correlation with literacy scores for both the receptive and expressive tests (all $r_s > .31$, all $p_s < .05$).

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Table 5

Spearman's Rho Correlations between the Tablet Test Count of Correct Answers for Expressive and Receptive Measures for Age and SES (N = 45)

Literacy measure	Age	SES
<i>Expressive count of correct scores</i>		
Upper case letter name	.32*	.39**
Lower case letter name	.33*	.40**
<i>Receptive formula score</i>		
Upper case letter name	.33*	.36*
Lower case letter name	.31*	.50**

Notes: SES = socioeconomic status. * $p < .05$; ** $p < .01$.

An independent samples t -test was used to compare test scores for upper and lower case letter knowledge between boys ($N = 29$) and girls ($N = 16$), using the count of correct responses for the expressive tablet tests, and the formula score for receptive tablet tests. No significant gender differences were found for any of the six tests (all t s < 0.14 , all p s $> .05$).

Correlations were used to further examine the relationships among the different tablet tests. Tablet formula scores were used for the receptive tests as these had demonstrated greater reliability and validity than the tablet correct count of answers for receptive tests. When tested with simple Spearman's rho correlations, strong relationships were found between all expressive correct count and receptive formula scores for upper and lower case letter name tests (all p s $< .01$) (see Table 6, top diagonal). Because age was shown to have a significant relationship with performance on literacy tests (see Table 5), partial Spearman's rho correlation analyses were conducted, with age as the control variable. Expressive correct count and receptive formula scores for upper and lower case letter name tests again showed strong associations (see Table 6, bottom diagonal).

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Table 6

Bivariate and Partial Spearman's Rho Correlations between Tablet Correct Count of Answers Scores and Tablet Formula Scores across Measures (N = 45)

Measure	Expressive upper case letter name	Expressive lower case letter name	Receptive upper case letter name	Receptive lower case letter name
Expressive upper case letter name	-	.97**	.88**	.88**
Expressive lower case letter name	.97**	-	.87**	.86**
Receptive upper case letter name	.86**	.85**	-	.90**
Receptive lower case letter name	.87**	.85**	.90**	-

Note. Top diagonal represents the simple correlations and the bottom diagonal represents the partial correlations using age as the control variable. All correlations are statistically significant at $p < .01$. * $p < .05$; ** $p < .01$.

Discussion

While educational apps are being increasingly used in classrooms in a range of curriculum areas, most of these applications have been used to support learning and teaching rather than assessment. The present study is the first study to our knowledge to report on and compare the validity and reliability of a tablet-based expressive and receptive test of early literacy skills in preschool children. The present study with its automated scoring software and touch-based operational features also extends and supports the work of previous research that has examined the use of computer based assessments for use in the early years classroom where teachers were required to narrate test questions and manually record and enter scores (Carson et al., 2011; Carson et al., 2015). The present work has also extended Neumann and Neumann's (2018) study and has shown that compared to traditional paper-pencil tests, the

EXPRESSIVE AND RECEPTIVE TABLET TESTS

assessment app test formats used in the present study was appropriate and user-friendly for young children. Moreover, tests of internal consistency, test-retest reliability, convergent validity and predictive validity indicated that, overall, the tablet expressive assessments were valid, reliable, and comparable to tablet receptive literacy and traditional literacy measures. More specifically, it shows that tablet devices installed with early literacy assessment apps with expressive or receptive response formats can be used by teachers and researchers to assess children's alphabet knowledge in the classroom.

Validity and reliability of tablet tests

Similar to the previous study by Neumann and Neumann (2018), two receptive test scoring approaches were examined: a count of correct answers and a correction for guessing formula. The internal consistency of both the measures for the present study was found to be excellent, which suggests that receptive tablet upper and lower case letter name tests are reliable measures. The receptive upper and lower case letter name tablet measures showed an excellent degree of agreement with paper-based expressive assessment measures. This suggests that the tablet tests might be used in the place of or in addition to the equivalent paper-pencil tests, thereby adding to assessment options available to teachers. Test-retest reliability was strong for the tablet receptive count of correct responses for upper and lower case letter name knowledge, indicating stability of the measure over time.

The reliability estimates were similar for the receptive test count of correct answers and the tablet formula score. Neumann and Neumann (2018) found consistency to be slightly greater for the formula score. Test-retest reliability was not as strong for the receptive tests in comparison to the expressive tests. Guessing by children was more likely to occur with four options to choose from when they are unsure of the correct answer. However, children who are guessing will not necessarily choose the same option when they guess from one test time to another. The greater opportunity to guess the correct answer with receptive tests may mean

EXPRESSIVE AND RECEPTIVE TABLET TESTS

that some scores were overinflated at Time 1 or Time 2, thus reducing the agreement between the two testing times.

The mean scores for the tablet receptive count score and formula score tests were found to be higher than those for the paper-pencil expressive tests. Thus, even after correcting for guessing, children performed better on the tablet-based receptive test than the paper-based expressive tests. One possible explanation for this is that receptive tests may not be measuring the exact same underlying cognitive process as expressive tests. Receptive-based assessment is more closely aligned with a recognition memory test, whereas an expressive literacy assessment is more aligned with a free recall memory test. In the receptive test, children were thus required to demonstrate letter name knowledge by matching cues together. One cue is the vocalization (e.g., “touch the letter P”) and the other cue is the letter form. In the expressive test, children were required to self-generate the vocalization from the letter form cue. This is not only a more difficult task but is also a task that requires some different underlying cognitive process with previous work highlighting how expressive tests are more cognitively demanding than receptive tests (Drouin et al. 2012). Further research is needed to examine such differences in expressive and receptive assessment tasks (Neumann & Neumann, 2018)

Another possible explanation for the receptive-expressive difference could be related to the mean age of the current sample ($M = 4.65$ years). Clark (1993) states that before a child can acquire competence in language production, they must first have competence in language comprehension. Furthermore, according to Girolametto, Wiigs, Smyth, Weitzman, and Pearce (2001), many young children experience delays with expressive language at age four years but have caught up to their peers by age five years. Such features of language development will result in lower test scores for expressive literacy assessments than for receptive literacy assessments. Longitudinal research is needed to examine developmental

EXPRESSIVE AND RECEPTIVE TABLET TESTS

differences associated with performance on tablet-based receptive and expressive letter name tests.

The expressive tablet assessment measure showed good reliability and validity, as demonstrated by its level of internal consistency, agreement with paper-based expressive alphabet letter tests (convergent validity), and test-retest reliability. The measures of all the tablet expressive tests showed excellent internal consistency, indicating that these tests are reliable measures of the constructs of upper and lower case name knowledge. The expressive tablet tests demonstrated no significant difference to their paper-based expressive counterparts. These findings confirm the construct and convergent validity of these tablet measures when testing expressive letter name knowledge.

The present study has not only extended the validation findings of a receptive tablet test (Neumann & Neumann, 2018) but has also validated a tablet expressive test and opened up new and broader questions about when and why each type of test (expressive or receptive) should be used in the classroom and how these varied formats could potentially benefit students with diverse needs. For example, a receptive tablet test that only requires tapping a screen with a finger and no verbal response may be appropriate for a student with a speech impairment or for testing multiple children at the same time. Furthermore, another avenue of investigation is how tablet receptive and expressive tests could be utilized not only to track student progress but also to drive student learning and mastery through repeated practice of assessment tasks.

Practical implications

Tablet devices connected to cloud-based data systems offer teachers a highly advantageous method of formatively assessing young children's letter and word knowledge (Bengfort, 2016; Carson et al., 2011). The present study has shown that tablet technology is a valid and reliable way to assess early literacy skills and importantly has shown that

EXPRESSIVE AND RECEPTIVE TABLET TESTS

expressive response formats can be used to test letter knowledge. The assessment app has scope for deriving unique measurements that are difficult to achieve with paper-based assessments. For example, the tablet assessments could be further developed to identify when errors are caused by feature similarities in alphabet letters. In the current study, it was clear to see from the digital test score dashboard when a child was confusing letters based on their shape or rotation (e.g., *m* with *w*), orientation (e.g., *d* with *b*), sound (e.g., *g* with *j*), or similarity to numbers (e.g., *l* with *1*). This ability to compare incorrect responses with the correct answers across the 26 letters could prove an invaluable and efficient tool for teachers. For example, such a digitally scored assessment would enable the immediate targeting of individual student letter shape confusions and the planning of specific classroom activities on learning to discriminate similar letters.

The tablet assessment apps also have the potential for making letter testing more dynamic thereby using a shortened number of test items based upon an individual's responses. For example, in the receptive test, a set of four letters (e.g., A, B, C, D) may be presented first. The narrator voice of the app asks the child to "Touch the letter B". The child touches the letter D. This response gives information that the child not only does not know B, but also that they do not know D. It may mean that it is not necessary to test B or D again. The next four options might be A, C, E, F. If they touch E when asked for F then new information is obtained that the child does not know these two letters either. With the future development of more advanced and enhanced tablet assessment software it might be possible to create dynamic random sequences to test every letter in significantly fewer than 26 trials.

An advantage of the tablet app used in the present study was that it was intuitive and user-friendly for children to operate independently. This was due the tablet's touchscreen interface, use of a finger to follow prompts from the narrator voice (e.g., "Touch the frog."), and easy to tap test items from the options provided. Compared to using a computer and

EXPRESSIVE AND RECEPTIVE TABLET TESTS

keyboard, this ease of tablet use offers distinct advantages for young children whose fine motor skills are still developing (Neumann & Neumann, 2014a). During the assessment sessions in the present study, children were observed to look forward to using the tablet and as a mobile device. The hand-held tablet was also found to be a flexible tool that can easily be used in various spaces of the classroom.

Limitations and future research

Due to the limited sample size of this study the findings cannot be generalized to all children. Further work is needed to validate expressive and receptive tablet apps with children from diverse backgrounds, ages, and varied abilities. For online assessments, the availability of reliable internet is also essential for efficient data collection. For example, if internet connectivity fails, test data can be lost, and children need to be retested. The speech recognition software also requires further refinement to ensure voice recognition data capture is accurate. For example, software voice-interpretation errors included *elf* for the letter *l* and *daddy yo* for *w*. In the present study, a voice recorder was used in addition to the tablet recordings to ensure data scoring accuracy. Indeed, previous researchers have commented on the difficulty of speech recognition to understand students' expressive language (see Baker, 2016). Further development of speech recognition software is required to ensure that young children's idiosyncratic ways of pronouncing letters and words are accurately distinguished.

Evaluating the online tablet assessments with various curriculum content in a wide age range of children and adults will help determine the benefits and disadvantages of this technology. Such an approach would provide a broader view on the practical benefits of tablets and assessment apps for classroom use and formative and summative assessment purposes. In addition to examining the use of expressive versus receptive response modalities in assessment apps, it is important to explore other design features such as multimodal features, interactive features, item presentation, operational features such as tap, drag and

EXPRESSIVE AND RECEPTIVE TABLET TESTS

drop, swipe, and presentation of animations for motivation and engagement (e.g., Hillman & Marshall, 2009; Marsh, 2016; Miller & Kocurek, 2017). This approach is needed for the development of optimal design features of assessment apps for use in the classroom to support the individual needs of diverse learners across curriculum areas. Finally, conducting ecological studies of teacher and student use of assessment apps in the classroom setting is essential to determine benefits of new technologies for test administration, scoring, reporting, and curriculum planning.

Conclusion

The present study has shown that using a tablet assessment with an expressive and receptive response format provides a valid and reliable measure of children's early literacy skills. These measures align well with traditional paper-pencil assessments. Providing teachers with the option of implementing either a receptive or expressive tablet test using a tablet or in paper form allows greater flexibility for delivering assessments that meet the needs of individual students. For example, children who have limited oral language skills or speech difficulties may benefit from using a receptive tablet-based assessment rather than an expressive test. The online data collection feature of the tablet assessment has the potential to increase the efficiency of collecting, storing and analyzing student data compared to traditional paper-based assessment methods. Further work is needed to examine the effects of other design features of apps to optimize children's test experiences and also the ecological validity of using online tablet-based assessment in the classroom to effectively link assessment outcomes with learning and teaching practices.

References

- Baker, E. (2016). Apps, iPads, and literacy: Examining the feasibility of speech recognition in a first-grade classroom. *Reading Research Quarterly*, 52, 291–310.
doi:10.1002/rrq.170
- Bengfort, J. (2016). Cloud in the classroom is a secure, efficient solution. *EdTech*. Retrieved from: <https://edtechmagazine.com/k12/article/2016/09/cloud-classroom-secure-efficient-solution>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18, 5–25. doi:10.1080/0969594x.2010.513678
- Broadfoot, P., & Black, P. (2004). Redefining assessment? Then first ten years of Assessment in Education. *Assessment in Education: Principles, Policy & Practice*, 11, 7–26.
doi:10.1080/0969594042000208976
- Carson, K. L. (2017). Reliability and predictive validity of preschool web-based phonological awareness assessment for identifying school-aged reading difficulty. *Communication Disorders Quarterly*, 39, 259–269. doi:10.1177/1525740116686166
- Carson, K., Boustead, T., & Gillon, G. (2015). Content validity to support the use of a computer-based phonological awareness screening and monitoring assessment (Com PASMA) in the classroom. *International Journal of Speech-Language Pathology*, 17, 500–510. doi:10.3109/17549507.2015.1016107
- Carson, K., Gillon, G., & Boustead, T. (2011). Computer-administrated versus paper-based assessment of school-entry phonological awareness ability. *Asia Pacific Journal of Speech, Language and Hearing*, 14, 85–101. doi: 10.1179/136132811805334876
- Clark, E. (1993). *The Lexicon in Acquisition* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511554377

EXPRESSIVE AND RECEPTIVE TABLET TESTS

- Crescenzi, L., Jewitt, C., & Price, S. (2014). The role of touch in preschool children's learning using iPad versus paper interaction. *Australian Journal of Language and Literacy*, 37, 87-95.
- Drouin, M., Horner, S. L., & Sondergeld, T. A. (2012). Alphabet knowledge in preschool: A Rasch model analysis. *Early Childhood Research Quarterly*, 27, 543–554. doi:10.1016/j.ecresq.2011.12.008
- Field, A. P. (2009). *Discovering statistics using SPSS*. London, England: SAGE.
- Fletcher-Watson, B. (2013). Apps for babies: implications for practice and policy. In: *ESRC Research Capacity Building Clusters: National Summit Conference 2013*. Aston Business School, Aston University, pp. 58-65 , ESRC Research Capacity Building Clusters; National Summit Conference 2013 , United Kingdom , 25-26 June.
- Flewitt, R., Messer, D., & Kucirkova, N. (2015). New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy*, 15, 289-310. doi: 1468798414533560.
- Frary, R. B. (1988). Formula scoring of multiple-choice tests (correct for guessing). *Educational Measurement Issues and Practice*, 7, 33–38. doi:10.1111/j.1745-3992.1988.tb00434.x
- Geist, E. (2014). Using tablet computers with toddlers and young pre-schoolers. *Young Children*, 69, 58-63.
- Girolametto, L., Wiigs, M., Smyth, R., Weitzman, E., & Pearce, P. S. (2001). Children with a history of expressive vocabulary delay: Outcomes at 5 years of age. *American Journal of Speech-Language Pathology*, 10, 358-369. doi:10.1044/1058-0360(2001/030)

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Goodwin, K., & Highfield, K. (2012). iTouch and iLearn: an examination of 'educational' Apps. In: *Early Education and Technology for Children conference* March 14-16, 2012, Salt Lake City, Utah.

Heritage, M. (2018). Assessment for learning as support for student self-regulation. *The Australian Educational Researcher*, 45, 51-63. <https://doi.org/10.1007/s13384-018-0261-3>

Herodotou, C. (2018). Young children and tablets: A systematic review of effects on learning and development. *Journal of Computer Assisted Learning*, 34, 1-9. <https://doi.org/10.1111/jcal.12220>

Hillman, M., & Marshall, J. (2009). Evaluation of digital media for emergent literacy. *Computers in the Schools*, 26, 256-270.

Hodges, D., Eames, C., & Coll, R. K. (2014). Theoretical perspectives on assessment in cooperative education placements. *Asia-Pacific Journal of Cooperative Education*, 15, 189–207.

Hollingshead, A. B. (1975). The four-factor index of social status. New Haven, CT: Unpublished manuscript. Yale University.

Koomen, M., & Zoanetti, N. (2016). Strategic planning tools for large-scale technology-based assessments. *Assessment in Education: Principles, Policy & Practice*, 25, 200–223. doi:10.1080/0969594x.2016.1173013

Lee, L. (2015) Digital media and young children's learning: A case study of using iPads in American preschools. *International Journal of Information and Education Technology*, 5, 947-950.

EXPRESSIVE AND RECEPTIVE TABLET TESTS

- Lesage, E., Valcke, M., & Sabbe, E. (2013). Scoring methods for multiple choice assessment in higher education – Is it still a matter of number right scoring or negative marking? *Studies in Educational Evaluation, 39*, 188–193. doi:10.1016/j.stueduc.2013.07.001
- Li, T. K., & Koo, M. Y. L. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155–163. doi:10.1016/j.jcm.2016.02.012
- Looney, A., Cumming, J., van der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy and Practice, 25*, 442-467.
<https://doi.org/10.1080/0969594X.2016.1268090>
- Lu, Y., Ottenbreit-Leftwich, A. T., Ding, Ai-Chu., & Glazewski, K. (2017). Experienced iPad-using early childhood teachers: Practices in the one-to-one iPad classroom. *Computers in the Schools, 34*, 1-15. <https://doi.org/10.1080/07380569.2017.1287543>
- Marsh, J. (2016). The digital literacy skills and competencies of children of pre-school age. *Media Education Studies and Social Research, 7*, 197-214.
- Marsh, J., Yamada-Rice, D., Bishop, J., Lahmar, J., Scott, F., Plowman, L., Piras, M., French, K., Robinson, P., Davis, S., Davenport, A., Winter, P. (2015). *Exploring Play and Creativity in Pre-Schoolers' Use of Apps: Technology and Play*. Economic and Social Research Council. Retrieved from: <http://www.techandplay.org/tap-media-pack.pdf>
- Miller, J. L., & Kocurek, C. A. (2017). Principles for educational game development for young children. *Journal of Children and Media, 11*, 314-329. doi:
[org/10.1080/17482798.2017.1308398](https://doi.org/10.1080/17482798.2017.1308398)
- Mohamad, A. J., Lakulu, M., & Samsudin, K. (2016). The needs of a framework of mobile application for kindergarten early reading: A review and analysis of the literature.

International Journal of Academic Research and Social Science, 6, 225–230.

doi:10.6007/ijarbss/v6-i12/2488

Neumann, M. M. (2018). Using tablets and apps to enhance emergent literacy skills in young children. *Early Childhood Research Quarterly*, 42, 239-246.

doi.org/10.1016/j.ecresq.2017.10.006

Neumann, M. M. (2016). Emergent Literacy Assessment app. Retrieved from iTunes:

<https://goo.gl/m7lYg7>

Neumann, M. M., & Neumann, D. L. (2014a). Touch screen tablets and emergent literacy.

Early Childhood Education Journal, 42, 231–239. doi:10.1007/s10643-013-0608-3

Neumann, M. M. & Neumann, D. L. (2014b). A measure of emerging print knowledge in young children. *Early Child Development and Care*, 184, 1142-1159. doi:

10.1080/03004430.2013.853054.

Neumann, M. M., & Neumann, D. L. (2018). Validation of a touch screen tablet assessment of literacy skills and a comparison with a traditional paper-based assessment.

International Journal of Research & Method in Education.

doi:10.1080/1743727X.2018.1498078

Oakley, G., Wildy, H., & Berman, Y. (2018). Multimodal digital text creation using tablets and open-ended creative apps to improve the literacy learning of children in early childhood classrooms. *Journal of Early Childhood*

Literacy. <https://doi.org/10.1177/1468798418779171>

Piasta, S. B., Petscher, Y., & Justice, L. M. (2012). How many letters should preschoolers in public programs know? The diagnostic efficiency of various preschool letter-naming benchmarks for predicting first-grade literacy achievement. *Journal of Educational Psychology*, 104, 945–958. doi:10.1037/a0027757

EXPRESSIVE AND RECEPTIVE TABLET TESTS

- Ralph, R., & Petrina, S. (2018). Social learning with mobile devices in preschool classrooms. *European Journal of STEM Education*, 3, 1-15. doi: doi.org/10.20897/ejsteme/3872
- Raskind, M. H., & Higgins, E. L. (1999). Speaking to read: The effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia*, 49, 251-281. doi: https://doi.org/10.1007/s11881-999-0026-9
- Silseth, K., & Gilje, O. (2019). Multimodal composition and assessment: a sociocultural perspective. *Assessment in Education: Principles, Policy, & Practice*, 1, 26-42. https://doi.org/10.1080/0969594X.2017.1297292
- Streiner, D. L., & Kottner, J. (2014). Recommendations for reporting the results of studies of instrument and scale development and testing. *Journal of Advanced Nursing*, 70, 1970–1979. doi:10.1111/jan.12402
- Vatalaro, A., Culp, A. M., Hahs-Vaughn, D. L., & Barnes, A. C. (2018). A quasi-experiment examining expressive and receptive vocabulary knowledge of preschool Head Start children using mobile media apps. *Early Childhood Education*, 46, 451–466. doi:10.1007/s10643-017-0877-3
- Watson, P. F., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. *Theriogenology*, 73, 1167–1179. doi:10.1016/j.theriogenology.2010.01.003
- Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, 50, 1093–1109. doi:10.1044/1092-4388(2007/076)

EXPRESSIVE AND RECEPTIVE TABLET TESTS

Woloshyn, V. W., Bajovic, M., & Worden, M. M. (2017). Promoting student-centred learning using iPads in a grade 1 classroom: Using the digital didactic framework to deconstruct instruction. *Computers in the Schools, 34*, 152-167.

Xie, H., Peng, J., Qin, M., Huang, X., Tian, F., & Zhou, Z. (2018). Can touchscreen devices be used to facilitate young children's learning? A meta-analysis of touchscreen learning effect. *Frontiers in Psychology, 9*.

<http://dx.doi.org/10.3389/fpsyg.2018.02580>