Touch Screen Assessment of High-Risk Infants' Word Knowledge

Évaluation par écran tactile du vocabulaire d'enfants présentant un risque élevé d'avoir un trouble de la parole et du langage

Rachel Hahn Arkenberg
Sharon Christ
Amanda Seidl

KEYWORDS
ASSESSMENT
DEVELOPMENT
TOUCH-SCREEN
EARLY INTERVENTION
LANGUAGE

Rachel Hahn Arkenberg, Sharon Christ, and Amanda Seidl

Purdue University, West Lafayette, IN, USA

Editor:
Stefano Rezzonico

Editor-in-Chief:
David H. McFarland

Abstract

Early identification of speech and language disorders is a priority for the field of speech-language pathology. The Computerized Comprehension Task is a promising tool for early assessment of language, because it preferentially taps strong word-referent associations (Friend & Keplinger, 2003), but its concurrent and predictive validity have not been examined in infants at high risk for speech and language disorders. We present preliminary findings related to using this tool with high-risk infants. We recruited 11 high-risk infants (having two or more risk factors) and 11 matched peers (14–24 months) to complete tests of speech and language at two time points, 6 months apart. Performance on the Computerized Comprehension Task was significantly correlated with standardized language measures for all infants. A mixed-effects model with corrections for small sample size and missing data revealed that the Computerized Comprehension Task and a more standard comprehension assessment were statistically significantly associated with expressive language outcomes 6 months post-baseline. This study provides preliminary data that the Computerized Comprehension Task could be a useful tool for early assessment of high-risk children and warrants further investigation in this population.

Abrégé

L'identification précoce des troubles de la parole et du langage est une priorité en orthophonie. La *Computerized Comprehension Task* [tâche de compréhension informatisée] est un outil prometteur pour l'évaluation du langage en bas-âge, car elle cible les associations mot-référent qui sont fortes (Friend et Keplinger, 2003). Cependant, les validités concordante et prédictive de cet outil n'ont pas encore été examinées chez les enfants présentant un risque élevé d'avoir un trouble de la parole et du langage. Cet article présente les résultats préliminaires suivant l'utilisation de cet outil auprès d'enfants présentant un tel risque. Nous avons recruté 11 enfants âgés de 14 à 24 mois présentant un risque élevé d'avoir un trouble de la parole et du langage (présence d'au moins deux facteurs de risque) et 11 enfants à faible risque qui ont été appariés sur l'âge et sur le niveau d'éducation de la mère. La parole et le langage de ces enfants ont été évalués à deux reprises (les deux évaluations étaient espacées par une période de six mois). Les performances de tous les enfants à la *Computerized Comprehension Task* étaient fortement corrélées à celles des mesures standardisées du langage. Les résultats d'un modèle mixte (corrigé pour la petite taille de l'échantillon et les données manquantes) révèlent une association significative entre les scores obtenus à la *Computerized Comprehension Task* et à une évaluation standardisée de la compréhension lors de l'évaluation initiale et le score obtenu à une mesure de langage expressif six mois plus tard. Les données préliminaires de cette étude suggèrent que la *Computerized Comprehension Task* pourrait être un outil utile pour l'évaluation de jeunes enfants présentant un risque élevé d'avoir un trouble de la parole et du langage et que des études plus approfondies auprès de cette population seraient justifiées.

Assessment of the communication of children under 2 years is a priority for the field of speech-language pathology, because early assessment and intervention can have a significant impact on academic and personal growth (Aram & Hall, 1989; Curtis et al., 2019; Lewis et al., 2011; McCormack et al., 2010; McKean et al., 2017; Wallace et al., 2015). Early assessment and intervention are particularly important in children who are at high risk for speech and language disorders. One of the most well-documented sources of risk is genetic, that is, having a sibling who has a communication disorder such as developmental language disorder (DLD; also called specific language impairment) or speech sound disorder. While both DLD (prevalence of 9.92%; Norbury et al., 2016) and speech sound disorder (prevalence of 16% at 3 years; Pennington & Bishop, 2009) are not highly prevalent in the general population, the odds are 1.7 times higher for siblings of children diagnosed with these disorders (Rudolph, 2017). Odds for speech and language disorders are also increased for children born preterm. Specifically, children born extremely preterm (less than 28 weeks) are 10 times as likely to have DLD and 4.4 times as likely to have speech sound disorder (Wolke et al., 2008).

However, risk for speech and language disorders is multifactorial, and one risk factor may not be sufficient to classify a child as "high risk" (Ebbels et al., 2019; Rudolph, 2017). Often a combination of risk factors, such as genetic risk, preterm birth, and low socioeconomic status, together contribute to overall risk (Hoff, 2003; Lewis et al., 2006, 2011; Rudolph, 2017). While many studies have examined risk factors for a variety of speech and language disorders, they have varied widely in sample size, ages, factors studied, and analyses, which limits comparison (Harrison & McLeod, 2010; Wallace et al., 2015). However, in a recent meta-analysis (i.e., Rudolph, 2017), studies of risk factors were systematically reviewed for quality and outcome variables before the meta-analysis was completed. The top five weighted risk factors (defined by highest odds ratio) are examined in our study: maternal education, family history, birth order, biological sex, and prematurity. Clearly, not every high-risk child develops a speech or language disorder, but, given the high-risk status of these children and the impact of these disorders on life outcomes, it is a priority to assess speech and language early and effectively in this population.

However, early assessment has some key maturation-related and methodological challenges. For example, speech production assessment, such as assessment of the phonetic inventory, is complicated by the wide age range for the onset of speech and clear confounds with motor development. Language comprehension, on the other hand, may be a more reliable measure of early language, and may enable standardized assessment at a younger age, since perception precedes production and lacks confounds associated with the protracted development of motor skills (Bornstein & Haynes, 1998; Davis & MacNeilage, 1990). Current methods of assessing comprehension in clinical settings are often indirect assessments, which assess the whole continuum of weak to strong associations between words and referents. For example, on the popular indirect parent report measure, the MacArthur-Bates Communicative Development Inventories (Bruckner et al., 2007; Fenson et al., 1993), a parent may correctly report that their child knows the word *dog*, even if their child has a relatively weak word-referent association such as only using the word *dog* to refer to their neighbour's pet.

On the other hand, direct comprehension assessment methods require both lexical retrieval and hypothesis testing, which means that they investigate strong, decontextualized associations (Yu & Smith, 2012). These decontextualized, stable associations appear to better predict downstream language in samples of primarily typically developing children (Friend et al., 2018; Schmitt, 2014). However, the concurrent and predictive value of direct assessment of comprehension is relatively unknown in children who are known to be at high risk for speech and language disorders.

Since it is vitally important to assess these high-risk children early, it is a priority to evaluate if direct assessment methods are predictive in this population. However, many direct comprehension assessment methods, such as the Mullen Scales of Early Learning (MSEL) or eye-tracking assessments, require a high degree of training, significant time investment, and/or expensive equipment and laboratory access, which limits their feasibility in clinical settings. Therefore, not only must we determine if direct assessments are predictive in this high-risk population, but it is also important to determine if clinically feasible direct assessments are predictive, because they have potential to lead to earlier identification.

One direct assessment that is relatively quick and simple to administer, while maintaining strong validity in developing children, is the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003). The CCT is a forced-choice measure of vocabulary comprehension, reflecting decontextualized word-referent associations. The CCT is validated for typically developing, English-speaking infants with high test-retest reliability ($r = .76$, $p < .05$), convergent validity ($r = .64$, $p < .05$), and improved performance over time ($r = .61$, $p < .05$; Friend & Keplinger, 2008; Friend et al., 2012, 2019). It correlates concurrently with parent report and predictively with a language sample, and the CCT (given at 22 months) predicts language at age 3 (Friend et al., 2019). It also identifies language deficits 2 years earlier than other methods, with similar sensitivity and specificity to Language Factor score (Friend et al., 2019). However, the CCT has not yet been studied in high-risk infants. Since this assessment method holds great promise for early identification of language deficits, it is imperative that we determine if it maintains concurrent and predictive validity in a high-risk population.

Therefore, we asked (a) Is decontextualized vocabulary, measured by the CCT, correlated with standardized and non-standardized measures of speech and language concurrently and predictively? and (b) Is performance on the CCT associated with 6-month expressive language outcomes in high-risk infants? We hypothesized that (a) decontextualized vocabulary measured by the CCT will have concurrent and predictive validity, demonstrated by correlation with other measures of speech and language given at the same time point and over time and (b) performance on the CCT will have predictive validity, demonstrated by association with 6-month expressive language outcomes in high-risk infants.

Method

Participants

Eleven low-risk and 11 high-risk infants age 14–24 months participated in this longitudinal study (19/22 completed both Sessions 1 and 2). Though this is a small sample, power analysis revealed that with only $n = 4$, differences could be detected in these groups with .80 power. Participants were monolingual infants recruited from our lab's database, which is compiled from birth records of local children, referrals from community speech-language pathologists, parents of siblings in other speech/language studies, and by searching a university-wide database of research participants. Infants could be defined as high risk in two ways. First, if they had a family history (i.e., a sibling with a reported speech/language problem), which is a well-established risk factor in the literature, and at least one other risk factor: second or later birth order, male biological sex, and/or preterm birth (< 37 weeks). Second, they were considered high risk if they had all three of the remaining risks: second or later birth order, male biological sex, and preterm birth (< 37 weeks). Each of these four risks have odds ratios of greater than 1.49 in well-controlled meta-analysis (Rudolph, 2017).

Socioeconomic status is another well-documented risk factor for language onset, but we were unsuccessful in recruiting a representative sample of this population. So instead of using it to categorize risk, we controlled for this by matching for maternal education, a proxy for socioeconomic status. The high- and low-risk groups were matched for both maternal education and age.

Procedure

During all visits, parents and infants were seen by a trained and supervised master's level speech-language pathology student researcher in a private clinic room with minimal distractions. All procedures were approved by Purdue University's Institutional Review Board (1610018380, 12/15/16). At the first visit, parents were introduced to the study and provided consent and case history information, and screenings were completed to determine eligibility. To be eligible for the study, infants needed to pass a hearing screening using otoacoustic emissions (Otoport OAE, Otodynamics) and demonstrate fine motor skills within normal limits (measured with the Fine Motor subtest of the MSEL). If the infants passed these two screenings, initial testing was completed.

Testing started with a production task to assess consonants present in the child's phonetic inventory. Each target word was elicited with three different toy exemplars and responses were audio-recorded (Shure PGXD1 Bodypack Transmitter and Shure PGXD4 Wireless Receiver). Words for this production task were selected to reflect a variety of consonants across word positions and word structures (see **Table 1**), based on production norms from WordBank (Frank et al., 2017). To elicit words, toys were taken out of a bag and described by saying, "This is a ____," and child repetitions of consonants in target words were transcribed phonetically live. All words were again transcribed for analysis using recordings, and consonants were considered present in the child's inventory if they appeared at least once.

Participants also completed two more scales from the MSEL related to language development (i.e., receptive language and expressive language; Mullen, 1995). The MSEL is a valid and reliable standardized test measuring cognitive ability and motor development. Each section takes approximately 10–20 minutes to complete. The MSEL was chosen because, like the CCT, it is a test that depends on infants' responses, as compared to popular measures solely based on parent report.

The final component of the session was the CCT. The CCT is a forced-choice direct assessment of language comprehension. It is delivered via touchscreen, but the touchscreen program must be administered by a live researcher/clinician, and the assessment takes about 5–10 minutes to administer. The experimenter gives verbal prompts embedded with the target word in child-directed speech, using the same prompt for each word class, such as "Where is the ____?", for nouns (for a detailed breakdown on word classifications, difficulty, and randomization, see Friend & Keplinger, 2003). Visual stimuli for the CCT

Table 1

Target Phonemes and Word Frames, Each Targeted With Three Unique Toy Exemplars

Object	Ball	Dog	Cat	Sock	Keys	Banana	Fish	Apple	Duck
Phonemes assessed	/b/ /l/	/d/ /g/	/k/ /t/	/s/ /k/	/k/ /z/	/b/ /n/	/f/ /ʃ/	/p/ /l/	/d/ /k/

include 41 pairs of high-quality digital images on a solid blue screen. Images are prototypical referents for nouns, verbs, and adjectives, controlled for salience to childhood. When the child touches the correct item, the program plays a reinforcer. This digital reinforcer includes an auditory stimulus of the lexical item produced in child-directed speech and a reinforcing sound such as a recording of the word *ball* with a bouncing sound. The sound is presented only when the infant touches the correct target as a motivator for engagement with the task. Infant sensitive screen areas encompass less than 50% of the screen area, so random touches have a low probability of being counted as correct.

To introduce this touchscreen task, five training trials were completed. During training trials, the examiner gave the child specific directions, modeled screen touches, and used a hand-over-hand technique to introduce the touch screen if the child did not touch when prompted. All 22 children completed all trials of the CCT, according to standardized administration procedures including standardized prompts as described in Friend and Keplinger (2003). Scores were both recorded manually by the examiner and automatically within the program, with 100% agreement.

Six months later, 9 high-risk and 10 low-risk participants returned and again completed the CCT and the two language subscales of the MSEL (attrition of 3).

Analysis

All tests were scored according to their prescribed methods. The phonetic inventory was scored by number of consonants present, the CCT was scored by number correct, and the subtests of the MSEL were scored according to the test's specifications. However, raw scores from the MSEL were utilized for analysis and comparison, since the phonetic inventory and CCT do not have standard scores, and age was matched in our sample. We used descriptive statistics, *t* tests, and mixed-effect modeling to test hypotheses in this study. Data were screened for missing information and outliers prior to analysis, with no outliers found.

To address our first question aimed at examining concurrent validity, correlations were calculated for scores on the CCT, language subtests of the MSEL, and our measure of phonetic inventory. To address our second question related to association with later outcomes, mixed-effects models were used. Mixed-effects modeling requires fewer assumptions than Analysis of Variance, can account for missing data, and is equipped for clustering of repeated assessments within child through use of clustered sandwich estimator and random intercepts. Also, it is able to handle continuous independent variables without the need to dichotomize or parcellate. A mixed-effect model with restricted maximum likelihood estimation and a Kenward-Roger correction was estimated. Restricted maximum likelihood estimation with correction decreases bias associated with small sample size, prevents inflation of Type 1 error rates, and accounts for missing data (Chawla et al., 2014; McNeish, 2017).

Results

All 22 infants who were screened were eligible to participate in the study. Ages ranged from 14.64 to 23.55 months in the high-risk group and 14.00 to 23.85 months in the low-risk group (see **Table 2**). The high-risk group had seven boys, and the low-risk group had five boys. Two-sample *t* tests were used to examine baseline group differences at the initial visit. They revealed no statistically significant differences between risk groups in age, $t(20) = 0.07$, $p = .94$; mother's education (a proxy for socioeconomic status), $t(20) = 0.32$, $p = .76$; number of ear infections, $t(20) = 0.72$, $p = .47$; or fine motor skills measured by the MSEL Fine Motor subtest, $t(20) = 0.64$, $p = .53$.

Performance was scored on phonetic inventory (number present: $M = 8.50$, $SD = 2.89$, Range = 5–16), CCT (number correct: $M = 16.72$, $SD = 6.85$), and raw MSEL scores (MSEL Receptive: $M = 17.41$, $SD = 3.92$; MSEL Expressive: $M = 17.23$, $SD = 1.99$). As seen in **Figure 1**, both receptive measures had larger ranges than expressive (MSEL Expressive Range = 14–21; MSEL Receptive Range = 13–27; CCT Range = 5–26). The high-risk group had a lower mean score on the CCT and a higher standard deviation ($M = 20.70$, $SD = 10.28$) than the low-risk group ($M = 23.10$, $SD = 7.86$), but these differences did not reach statistical significance, $t(20) = 1.19$, $p = .25$. This is not unexpected, since not all "at risk" children will go on to develop speech or language disorders.

Table 2

Participant Characteristics for all Infants, Grouped by Risk Category

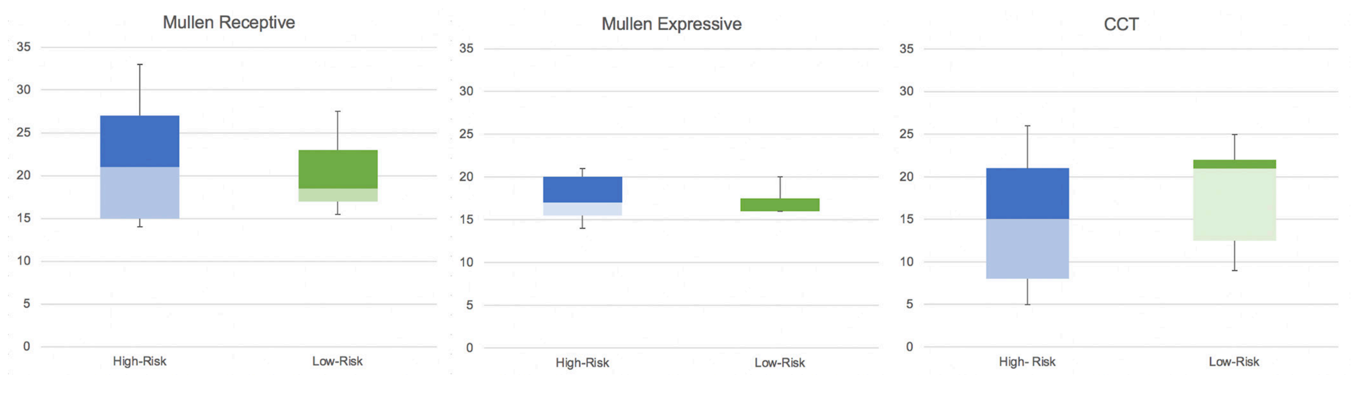
Age (months)	Mother's education (years)	Reason for risk	Number of risk factors
14.64	15	Sibling/parent with SSD, male, birth order	3
15.00	16	Sibling with SSD, male, birth order	3
15.03	16	Sibling with DLD/parents with SSD, male, birth order	3
15.66	12	Sibling with DLD/SSD, birth order	2
15.69	18	Sibling with SSD, male, birth order	3
16.12	16	Sibling with SSD, preterm, birth order	3
17.57	16	Sibling with DLD/SSD, male, birth order	3
18.39	16	Sibling with SSD, birth order	2
19.97	16	Preterm, male, birth order	3
23.50	19	Sibling/family with DLD/SSD, male, birth order	3
23.55	16	Sibling with DLD, birth order	2
14.00	18		0
14.74	12	Male	1
14.84	16		0
15.39	20	Male	1
15.69	18		0
17.11	16	Birth order	1
18.75	18	Male, Birth order	2
19.14	16		0
19.47	16	Male	1
23.32	13		0
23.85	16	Male, Birth order	2

Note. SSD = speech sound disorder; DLD = developmental language disorder.

To examine concurrent validity across tasks, correlations were calculated between measures for each time point (**Table 3**), which revealed large, significant correlations between each pair. Correlations were also computed for measures between Time 1 and Time 2 to investigate predictive validity. These also revealed large, significant correlations between each pair (**Table 4**).

To further investigate prediction of 6-month language outcomes, a mixed-effects regression model with Kenward–Roger corrections was estimated for prediction of performance on the expressive language subtest of the MSEL. Since there were multiple measures within subjects, a random intercept was included to capture inter-participant variability. We used backward step-wise selection to sequentially remove non-significant predictors: CCT score, age, risk, score on receptive subtest of the MSEL, and interaction between risk and CCT. There was no interaction, and CCT and score on the receptive subtest of the MSEL were the only significant predictors. All possible models from significant predictors were contrasted using Akaike

Figure 1



Performance on Mullen Scales of Early Learning (MSEL) and Computerized Comprehension Task (CCT) at Time 1, separated by risk group.

Table 3

Concurrent Validity of Speech and Language Measures at Time 1 and at Time 2

	Time 1 (n = 22)		Time 2 (n = 19)	
	Mullen Receptive	Mullen Expressive	Mullen Receptive	Mullen Expressive
Mullen Receptive	-			
Mullen Expressive	.755***	-	.606**	-
CCT	.771***	.752***	.674**	.650**

Note. CCT = Computerized Comprehension Task.
** $p < .01$ (2-tailed). *** $p < .001$ (2-tailed).

Table 4

Predictive Validity of Speech and Language Measures at Time 1 and at Time 2 (n = 19)

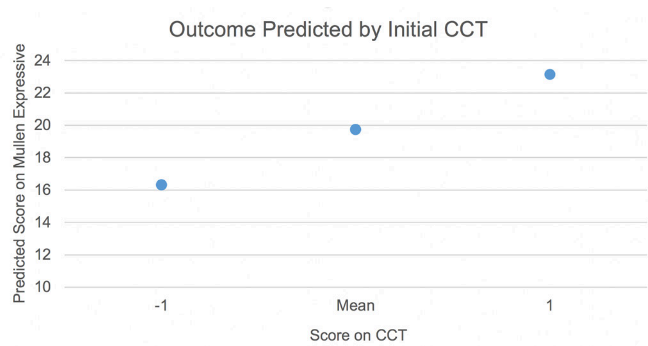
	T1 Mullen Receptive	T1 Mullen Expressive	T1 CCT
T2 Mullen Receptive	.532*	.570*	.466*
T2 Mullen Expressive	.776***	.807***	778***
T2 CCT	.598**	.628**	.610**

Note. T1 = Time 1; T2 = Time 2; CCT= Computerized Comprehension Task.
* $p < .05$ (2-tailed). ** $p < .01$ (2-tailed). *** $p < .001$ (2-tailed).

Information Criterion (Posada & Buckley, 2004) to identify the most parsimonious model with the best fit. The lowest Akaike Information Criterion value (i.e., 212.59) was obtained for CCT and MSEL receptive scores; however, all three models were within one point of this value, so they cannot fully be distinguished in terms of fit.

Selecting the model with the lowest Akaike Information Criterion, the final model, $F(2, 36.83) = 29.32$, $p < .001$, included CCT, $B = 0.21$, $t = 2.26$, $p < .03$, and Mullen Receptive score, $B = 0.34$, $t = 2.17$, $p < .04$. Proportion of the variance explained at the level of the person was 1.6×10^{18} and proportion of variance explained at the level of time was 1.4998. As seen in **Figure 2**, infants who score higher on the CCT, regardless of risk, score higher on expressive language 6 months later. Also, there is about a 7-point difference between model estimated expressive language scores for infants one standard deviation below the mean on the CCT and one standard deviation above the mean on the CCT 6 months prior.

Figure 2



Six-month language outcome on the Mullen Scales of Early Learning (MSEL) Expressive Language Subtest as a function of Computerized Comprehension Task (CCT) score (values = 1 *SD* below the mean, mean, 1 *SD* above the mean).

Discussion

In this study, we sought to preliminarily investigate the concurrent and predictive validity of a clinically feasible direct assessment of comprehension in high-risk infants and their low-risk peers. Receptive language has been established as an effective means to assess early language abilities and is associated with later language outcomes, likely because it captures variability at an earlier time than tests of expressive language (Bornstein & Haynes, 1998; Feldman et al., 2000; Friend et al., 2012, 2019; Kuhl, 2009; Marchman & Fernald, 2008; Tsao et al., 2004). More specifically, direct assessment is reported to tap into infants' strong word-referent associations, which has more predictive power than indirect assessment (Friend et al., 2019).

Indirect assessments, like parental report measures, are hypothesized to reflect the full range of strong to weak word-referent associations, whereas decontextualized direct assessment preferentially gives insight to strong associations. Additionally, there is some bias associated with parent report (Feldman et al., 2000; Oliver et al., 2002). Our study provides preliminary evidence that direct assessment of comprehension is valid in high-risk infants. Correlations and mixed models revealed good concurrent and predictive validity of the CCT in this population (Post, 2016).

Our results revealed that CCT scores in our two groups (high-, low-risk) were not statistically significantly different, but that the CCT was a feasible task for both groups. Importantly, overall scores on the CCT were highly correlated with a standardized, well-established productive test of receptive language ($r = .77$, $p < .001$) and with a typical, non-standardized phonetic inventory ($r = .75$, $p < .001$) for both low-risk and high-risk infants. Additionally, a mixed-effects model revealed that the CCT—not risk group or the interaction between risk and CCT—was significantly associated with expressive language, measured 6 months later.

Not only was the CCT highly associated with more established tests both concurrently and predictively, the CCT has an advantage over these measures since it is significantly faster, less expensive, and requires fewer manipulatives than the MSEL or eye-tracking assessment. Since this preliminary study had a small sample size, did not include blinding, had a wide range of ages, and had a limited longitudinal follow-up, its results (in isolation) should be interpreted with caution. However, when taken in the context of previous studies on the CCT completed by unrelated laboratories, it adds support to the validity of the CCT for direct measurement of receptive language in infants (Friend et al., 2019), and it provides initial data that supports further study of the CCT in high-risk infants.

Conclusion

This preliminary study revealed that the CCT has good concurrent and predictive validity in a sample of high-risk infants and matched low-risk peers. Future studies are warranted, and large-scale longitudinal data exploring both language and academic outcomes in low-risk and high-risk infants is needed to support this preliminary sample and to provide standardized scores for clinical utility.

References

Aram, D. M., & Hall, N. E. (1989). Longitudinal follow-up of children with preschool communication disorders: Treatment implications. *School Psychology Review, 18*(4), 487–501. https://doi.org/10.1080/02796015.1989.12085444

Bornstein, M. H., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development, 69*(3), 654–671. https://doi.org/10.1111/j.1467-8624.1998.tb06235

Bruckner, C., Yoder, P., Stone, W., & Saylor, M. (2007). Construct validity of the MCDI-I Receptive Vocabulary scale can be improved: Differential item functioning between toddlers with autism spectrum disorders and typically developing infants. *Journal of Speech, Language, and Hearing Research, 50*(6), 1631–1638. https://doi.org/10.1044/1092-4388(2007/110)

Chawla, A., Maiti, T., & Sinha, S. (2014). *Kenward-Roger approximation for linear mixed models with missing covariates* (Technical report). Michigan State University.

Curtis, P. R., Kaiser, A. P., Estabrook, R., & Roberts, M. Y. (2019). The longitudinal effects of early language intervention on children's problem behaviors. *Child Development, 90*(2), 576–592. https://doi.org/10.1111/cdev.12942

Davis, B. L., & MacNeilage, P. F. (1990). Acquisition of correct vowel production. *Journal of Speech, Language, and Hearing Research, 33*(1), 16–27. https://doi.org/10.1044/jshr.3301.16

Ebbels, S. H., McCartney, E., Slonims, V., Dockrell, J. E., & Norbury, C. F. (2019). Evidence-based pathways to intervention for children with language disorders. *International Journal of Language & Communication Disorders, 54*(1), 3–19. https://doi.org/10.1111/1460-6984.12387

[1]A model was also run where random effects were added for inter-item variability on the CCT, but this revealed no meaningful differences, so it is not reported here.

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development, 71*(2), 310–322. https://doi.org/10.1111/1467-8624.00146

Fenson, L., Dale, P. S., Reznik, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual.* Singular Publishing Group.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data*. *Journal of Child Language, 44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers, 35*(2), 302–309. https://doi.org/10.3758/BF03202556

Friend, M., & Keplinger, M. (2008). Reliability and validity of the Computerized Comprehension Task (CCT): Data from American English and Mexican Spanish infants*. *Journal of Child Language, 35*(1), 77–98. https://doi.org/10.1017/S0305000907008264

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental Psychology, 48*(1), 136–148. https://doi.org/10.1037/a0025511

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018). "A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness": Correction to Friend et al. (2018). *Developmental Psychology, 54*(7), 1289–1289. https://doi.org/10.1037/dev0000590

Friend, M., Smolak, E., Patrucco-Nanchen, T., Poulin-Dubois, D., & Zesiger, P. (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. *Developmental Psychology, 55*(1), 9–22. https://doi.org/10.1037/dev0000617

Harrison, L. J., & McLeod, S. (2010). Risk and protective factors associated with speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *Journal of Speech, Language, and Hearing Research, 53*(2), 508–529. https://doi.org/10.1044/1092-4388(2009/08-0086)

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

Kuhl, P. K. (2009). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In J. Colombo, P. McCardle, & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research disorders* (pp. 213–244). Psychology Press.

Lewis, B. A., Avrich, A. A., Freebairn, L. A., Hansen, A. J., Sucheston, L. E., Kuo, I., Taylor, H. G., Iyengar, S. K., & Stein, C. M. (2011). Literacy outcomes of children with early childhood speech sound disorders: Impact of endophenotypes. *Journal of Speech, Language, and Hearing Research, 54*(6), 1628–1643. https://doi.org/10.1044/1092-4388(2011/10-0124)

Lewis, B. A., Shriberg, L. D., Freebairn, L. A., Hansen, A. J., Stein, C. M., Taylor, H. G., & Iyengar, S. K. (2006). The genetic bases of speech sound disorders: Evidence from spoken and written language. *Journal of Speech, Language, and Hearing Research, 49*(6), 1294–1312. https://doi.org/10.1044/1092-4388(2006/093)

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*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

McCormack, J., McLeod, S., Harrison, L. J., & McAllister, L. (2010). The impact of speech impairment in early childhood: Investigating parents' and speech-language pathologists' perspectives using the ICF-CY. *Journal of Communication Disorders, 43*(5), 378–396. https://doi.org/10.1016/j.jcomdis.2010.04.009

McKean, C., Reilly, S., Bavin, E. L., Bretherton, L., Cini, E., Conway, L., Cook, F., Eadie, P., Prior, M., Wake, M., & Mensah, F. (2017). Language outcomes at 7 years: Early predictors and co-occurring difficulties. *Pediatrics, 139*(3), Article e20161684. https://doi.org/10.1542/peds.2016-1684

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research, 52*(5), 661–670. https://doi.org/10.1080/00273171.2017.1344538

Mullen, E. (1995). *Mullen Scales of Early Learning.* Circle Pines.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *The Journal of Child Psychology and Psychiatry, 57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Oliver, B., Dale, P., Saudino, K., Petrill, S., Pike, A., & Plomin, R. (2002). The validity of a parent-based assessment of cognitive abilities in three-year olds*. *Early Child Development and Care, 172*(4), 337–348. https://doi.org/10.1080/03004430212713

Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*(1), 283–306. https://doi.org/10.1146/annurev.psych.60.110707.163548

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology, 53*(5), 793–808. https://doi.org/10.1080/10635150490522304

Post, M. W. (2016). What to do with "moderate" reliability and validity coefficients? *Archives of Physical Medicine and Rehabilitation, 97*(7), 1051–1052. https://doi.org/10.1016/j.apmr.2016.04.001

Rudolph, J. M. (2017). Case history risk factors for specific language impairment: A systematic review and meta-analysis. *American Journal of Speech-Language Pathology, 26*(3), 991–1010. https://doi.org/10.1044/2016_AJSLP-15-0181

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951. https://doi.org/10.1111/lang.12077

Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*(4), 1067–1084. https://doi.org/10.1111/j.1467-8624.2004.00726.x

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., Cullen, K., & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics, 136*(2), e448–e462. https://doi.org/10.1542/peds.2014-3889

Wolke, D., Samara, M., Bracewell, M., & Marlow, N. (2008). Specific language difficulties and school achievement in children born at 25 weeks of gestation or less. *The Journal of Pediatrics, 152*(2), 256–262.e1. https://doi.org/10.1016/j.jpeds.2007.06.043

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Authors' Note

Correspondence concerning this article should be addressed to Rachel Hahn Arkenberg, Department of SLHS, Purdue University, West Lafayette, IN, United States, 47907. Email: **hahnr@purdue.edu.**

Acknowledgments

We would like to acknowledge Margaret Friend for her valuable insight on this study, as well as express our gratitude to all the infants and families who participated in this study.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.