Examining Performance on a Process-Based Assessment of Word Learning in Relation to Vocabulary Knowledge and Learning in Vocabulary Intervention

Elizabeth Spencer Kelley, Ph.D., CCC-SLP[1] and Howard Goldstein, Ph.D., CCC-SLP[2]

ABSTRACT

Vocabulary knowledge of young children, as a well-established predictor of later reading comprehension, is an important domain for assessment and intervention. Standardized, knowledge-based measures are commonly used by speech-language pathologists (SLPs) to describe existing vocabulary knowledge and to provide comparisons to same-age peers. Process-based assessments of word learning can be helpful to provide information about how children may respond to learning opportunities and to inform treatment decisions. This article presents an exploratory study of the relation among vocabulary knowledge, word learning, and learning in vocabulary intervention in preschool children. The study examines the potential of a process-based assessment of word learning to predict response to vocabulary intervention. Participants completed a static, knowledge-based measure of vocabulary knowledge, a process-based assessment of word learning, and between 3 and 11 weeks of vocabulary intervention. Vocabulary knowledge, performance on the process-based assessment of word learning, and learning in vocabulary intervention were strongly related. SLPs might make use of the information provided by a process-based assessment of word learning to determine the appropriate intensity of intervention and to identify areas of phonological and semantic knowledge to target during intervention.

KEYWORDS: assessment, vocabulary, word learning, measurement, preschool

[1]Department of Speech, Language, and Hearing Sciences at the University of Missouri-Columbia, Columbia, Missouri; [2]Department of Communication Sciences and Disorders at the University of South Florida, Tampa, Florida.

Address for correspondence: Elizabeth Spencer Kelley, Ph.D., CCC-SLP, Department of Speech, Language, and Hearing Sciences at the University of Missouri-Columbia, 315 Lewis Hall, Columbia, MO 65211 (e-mail: kelleyej@missouri.edu).

Semin Speech Lang 2019;40:344–358.

DOI: https://doi.org/10.1055/s-0039-1688447.
ISSN 0734-0478.

Learning Outcomes: As a result of this activity, the reader will be able to (1) describe several approaches to the examination of vocabulary knowledge and word learning, including process-based assessments; (2) discuss the use of a process-based assessment of word learning to predict learning in response to vocabulary intervention; and (3) explain how information from a process-based assessment of word learning can inform treatment decisions.

Vocabulary knowledge in preschool is a well-established predictor of future reading comprehension abilities.[1,2] Children with limited vocabulary in preschool are at high-risk for future reading and other academic difficulties.[3] However, preschool experiences that improve vocabulary knowledge can foster later reading comprehension[4] meaning that early identification of children who would benefit from vocabulary intervention may prevent future academic difficulties. Commonly used measures of vocabulary knowledge, such as standardized, norm-referenced measures of single-word vocabulary, can describe the existing receptive and expressive vocabulary knowledge of young children. In a clinical context, such measures are useful to provide comparisons to same-age peers and can indicate a general need for vocabulary intervention. However, these knowledge-based, static measures of vocabulary knowledge are highly dependent on experience and, thus, are reflective of a child's word learning environment rather than an indication of a child's word learning abilities. To complement static measures of vocabulary knowledge, process-based assessments of word learning ability can provide valuable information to guide treatment decisions about the intensity and delivery of intervention.

Static measure of vocabulary knowledge. Norm-referenced assessments are often required to determine eligibility for special education services. However, commonly-used assessments are primarily static measures that assess knowledge or ability at a single point in time. For example, the Peabody Picture Vocabulary Test (PPVT[5]) a widely-used norm-referenced, static measure of receptive vocabulary, provides information about acquired vocabulary knowledge (i.e., the words a child knows at the time of assessment). Static assessments are poor measures of the learning process (i.e., how a child learns words) and are strongly influenced by a child's previous experiences. As an example, for children with diverse backgrounds and language learning experiences, a low score on the PPVT might be an indication of limited language learning experiences as opposed to limited language learning abilities.

Process-based assessment. In contrast to static, knowledge-based measures, a process-based assessment is designed to describe the learning process. Several techniques can be incorporated into a process-based assessment, such as hierarchical prompting or a test-teach-test paradigm.[6] In hierarchical, or graduated, prompting, children are provided with a predetermined set of prompts to identify the amount of support needed to reach a correct response. In the test-teach-test paradigm, often referred to as dynamic assessment, the teaching phase consists of supportive instruction designed to facilitate learning and the repeated testing phases include prompting or scoring that is sensitive to small changes in knowledge.

Process-based assessments of language have been used frequently with populations of culturally and linguistically diverse children,[7–9] in particular to discriminate children with and without language impairments.[10–12] Dynamic assessments of word learning, specifically, have been useful in distinguishing between children with and without language difficulties in bilingual preschoolers.[12] Dynamic assessments of word learning also have been used to identify children with language impairment within groups of children who have been referred for speech and language services.[13–15] Kelley[16] reported on the development of a process-based assessment of explicit word knowledge. In the explicit word learning measure (EWL), children were exposed to brief teaching trials of novel words, immediately followed by probes for production and definitional knowledge. Because stimuli were novel words, the need for an initial 'test' phase was eliminated. The measure was

repeated in four sessions across several weeks and hierarchical prompting and incremental scoring allowed for sensitive measurement of learning. In this preliminary investigation, performance on the EWL was correlated with scores on static, knowledge-based measures of vocabulary and word learning on an incidental task, suggesting that the measure has the potential to provide meaningful information about word learning in young children. Together, these findings indicate that process-based assessments of word learning have the potential to provide valuable information to inform clinical decisions.

Process-based measures to inform treatment decisions. For speech-language pathologists (SLPs), another potential application of process-based assessments of word learning may be to identify children who struggle to learn words. In particular, performance on a process-based assessment of word learning may help identify children who require intense, explicit intervention. Increasing intensity of vocabulary instruction can improve learning for children at-risk of language difficulties.[17,18] Indeed, word-learning deficits in children with developmental language disorders can be ameliorated by increasing the number of learning opportunities.[19,20]

Process-based assessments may be particularly useful with two groups of children served by SLPs, children with developmental language disorders (DLD) and children from low SES families. Studies of the word learning process have found that, as a group, children with DLD often perform poorly on measures of word learning relative to peers with typical language. Children with DLD comprehend and produce fewer words than peers with typical language,[21–26] require more trials to learn new words,[27] and appear to be less able to learn both labels and semantic features (e.g., color, speed).[28,29] However, in other studies, children with DLD have performed similarly to peers with typical language on word learning tasks.[30,31] A careful review of this literature indicates that word learning appears to be particularly difficult for some children with DLD.[21,22,27] For example, Kiernan and Gray[21] found that, eight of the 30 children in the DLD group produced fewer words than any of the children in the normal language group, although there was substantial overlap in children with and without DLD. However, scores on static, knowledge-based measures (e.g., PPVT) did not identify poor word learners. Scores for the poor word learners were within age expectations (e.g., standard scores 85–97[22]) and overlapped with scores of good word learners.[27] In these studies, the measure of word learning was useful in identifying poor word learners within groups of children with DLD.

Another population often served by SLPs is children from families with low socioeconomic status (SES). Because SES-related differences in linguistic input have a large effect on vocabulary knowledge,[32] many children from low SES families may have lower scores on static, knowledge-based measures of vocabulary knowledge relative to peers with higher SES.[33,34] However, there is little evidence that these scores are an indication of poor word learning abilities. Within a group of African-American kindergartners, Burton and Watkins[7] found that risk status, as determined by several socioeconomic indicators, was strongly related to PPVT scores, but not related to performance on a process-based assessment of word learning. Similarly, Horton-Ikard and Ellis Weismer[35] observed SES-related differences on the PPVT but not on a fast-mapping word learning task. Together, these findings suggest that, particularly within groups of children from families with low SES, SLPs might consider a process-based assessment of word learning as a way to distinguish between children who have limited vocabulary as the result of experience and children who have poor word learning proficiency.

Although process-based assessments can identify children with language impairment, little is known about how performance relates to learning in intervention. Performance on dynamic assessment of language is strongly related to language growth[36] and performance on a dynamic assessment of word learning predicts growth in vocabulary over the next six months.[15] These findings indicate that process-based approaches may be effective in predicting how children will respond to vocabulary intervention. An important next step is to examine the relation among measures of vocabulary knowledge, word learning, and learning in vocabulary intervention to understand the potential of a process-based

assessment of word learning to guide treatment decisions.

THE CURRENT STUDY

Although static, knowledge-based assessments of single word vocabulary are appropriate to measure the vocabulary knowledge of young children, they are not likely to be sensitive to differences in word learning ability. In contrast, process-based assessments of language and word learning have been useful in identifying children who have language impairments and who are poor word learners, respectively. Process-based assessments of word learning may have the potential to inform treatment decisions by helping SLPs match poor word learners to appropriately intense vocabulary interventions.

The purpose of the current study was to explore the relation between vocabulary knowledge, word learning proficiency, and response to vocabulary intervention. First, descriptive analyses were conducted to examine performance on the process-based measure of word learning (e.g., floor or ceiling effects). Next, correlations were conducted to examine the relation among performance on a static, knowledge-based measure of vocabulary, the process-based assessment of explicit word learning, and learning in the context of vocabulary intervention. The hypothesis was that performance on the process-based assessment of word learning would be strongly related to learning in vocabulary intervention. Finally, exploratory analyses were conducted to examine the ways in which a process-based assessment of word learning might inform treatment decisions by SLPs.

METHOD

All study procedures were approved by the University of Missouri's Institutional Review Board.

Participants

Participants were 16 preschool children (6 girls, 10 boys) between the ages of 45 and 62 months (mean age 54 months). Participants were recruited as part of another study to evaluate a new component of the *Story Friends* intervention, described below. To obtain a sample that represented a wide range of SES, 10 children were recruited from two Head Start centers that served families who met eligibility guidelines for low income and 6 children attended a private preschool that served families with middle and high SES. The goal of recruitment in the two different types of classrooms was to select a group of children with a wide range of vocabulary scores to inform the design of *Story Friends.* In all classrooms, parents of children in the classroom were invited to participate by their teachers using flyers and informed consent documents. Children who had scores below 70 on the PPVT or who spoke very little English per teacher report were excluded based on previous studies that indicated these children were unlikely to benefit from the *Story Friends* program. This excluded 13 children from the Head Start classrooms and no children from the private preschool classroom. To describe overall language abilities, children completed the core language scale of the Clinical Evaluation of Language Fundamentals Preschool-2 (CELF-P2).[37] To accommodate potential variations in the use of dialect, the CELF-P2 was scored using the dialectical options presented in the assessment manual. Due to time constraints for testing, only 13 of the 16 children completed the CELF-P2 (M= 90.92, SD = 18.66). Of these children, 7 had scores between 85 and 115, placing them in the average range of language abilities for children their age; 5 had scores below 85, placing them below average; and 1 had a score above 115, placing her above average. Additional demographic information was requested via a family survey, but because only six families returned surveys, no demographic information is reported here.

Vocabulary intervention. The *Story Friends* vocabulary program is an automated, explicit vocabulary intervention designed for preschool children. The program has been evaluated in previous studies[38–41] of preschool children, with large learning effects for vocabulary presented as part of the intervention (Cohen's $f^2 = 0.70$). Additional detail on the program is provided in previous publications.[42]

Children in the current study were part of a pilot study to evaluate a newly developed series of books, *The Ocean Friends.* Each book included embedded lessons for four challenging vocabulary

words (e.g., *curious, drowsy, discover, create*). Embedded lessons provided explicit instruction with a child-friendly definition, supportive contexts, and multiple opportunities to respond. In each week of intervention, children listened to the same book three times on different days and thus, received instruction on four target vocabulary words per week. A different book was presented each week. Research staff conducted small-group listening centers three to four days per week to ensure that each child listened three times to that week's book. In the current study, children completed between 3 and 11 books of the *Story Friends* intervention. The number of books completed by each child varied due to children leaving the center and to scheduling constraints of the preschool (i.e., summer break).

Measures

The primary variables for the current study were (a) vocabulary knowledge, measured by a static measure of single-word receptive vocabulary, the Peabody Picture Vocabulary Test (PPVT-IV)[5]; (b) word learning proficiency, measured by a process-based assessment of word learning, Explicit Word Learning (EWL)[16]; and (c) the learning that occurred in the Story Friends intervention,[43] as measured by a definition test of targeted vocabulary.

Description of the Explicit Word Learning Measure. The EWL is a process-based assessment designed to describe word learning proficiency of preschool children.[16] The EWL includes brief, explicit teaching trials for novel words followed by probes for definitional knowledge and production. The EWL was presented on tablet computers that displayed high-resolution photographs and a standard script read by the examiner.

The EWL test items included four target nonwords (e.g., *yame*). Nonwords were chosen to ensure that children did not have previous experiences with targets. Because nonwords were used, it was not necessary to pretest children on their knowledge of the stimuli. The nonwords had low-probability, high neighborhood density phonotactic patterns, identified from Storkel, Armbruster, and Hogan.[44] The nonwords were assigned synonyms likely to be familiar to preschool children (e.g., *yame* means *happy*). Familiarity was determined by consulting published databases to select words with early age of acquisition.[45]

EWL teaching trials. Teaching trials included brief, explicit instruction designed to facilitate learning of the non-word target by presenting consistent instructional language and including multiple opportunities for children to interact and respond. Teaching trials for all four words took ~5 minutes to deliver. For each word, the teaching trial provided frequent presentation of the word (20 times) and the definition (10 times). Words were presented with accompanying pictures that provided contextual information and verbal scripts including child-friendly contexts and examples. Opportunities to respond included prompts to say the target word (2 times; "*Yame. Say yame.*"), prompts to say the definition (2 times; "*Tell me, what does yame mean?*), and a prompt to use a gesture, facial expression, or other verbal response (1 time; "*Show me how you look when you feel yame.*"). Teaching trials for all four words were presented in a predetermined sequence in which children completed a teaching trial for one word before moving on to the next. This sequence was counterbalanced across children.

EWL probes. Probes for definitional knowledge and production were administered immediately following the teaching trials. Definitional probes were always administered first, to prevent carryover from the production probe to definitional responses, and words were assessed in the same random sequence as the teaching trial. Administration of each probe took ~3 minutes.

Definitional probe. On the definitional probe, children were asked to respond to one open-ended question, worth up to two points, and four yes/no questions, worth up to one additional point. The open-ended question asked for a definition (e.g., *What does yame mean?*). Children received 2 points for a complete, correct definition (e.g., "happy"), 1 point for a partial or related response (e.g., "when you smile"), or 0 points for an incorrect, 'I don't know', or no-response. The four yes/no questions included a pair of questions that assessed knowledge of the definition (e.g., *Does yame mean happy?*) and a pair of questions that assessed contextual knowledge of the word (e.g., *If you fell down and got hurt, would you*

feel yame?) Within each pair, the correct answer for one question was 'yes' and the correct answer for the other question was 'no'. To reduce the likelihood that credit was awarded due to chance, children who responded correctly to at least three of the four questions received 1 point total. The procedures for the yes/no questions were adapted from measures used in similar studies of vocabulary intervention.[46] Thus, children could receive up to three points per word on the definitional probe with a maximum score of 12 points on the measure (3 points each for 4 words). This incremental scoring was designed to capture a range of word knowledge, including partial knowledge of the words.

Production probe. The children's ability to produce the target words was assessed using a hierarchical prompting procedure adapted from Burton and Watkins.[7] On the production probe, children were shown a picture, different from the picture used in the teaching trial, and were presented with a series of prompts designed to elicit the target non-word. Prior to the first item, children were reminded to use their 'new' words. The first prompt was open ended (e.g., *Look at this picture. These boys feel…*). The second prompt included a semantic cue (e.g., *Can you tell me another word for happy?*). The third prompt provided a phonological cue (*e.g. Another word for happy is /y/…*). The fourth prompt gave an indirect model (e.g., *Another word for happy is yame. How do these boys feel? They feel …*). At each prompting level, if a child responded correctly, successive prompts were not delivered. Scores for the production probe were assigned by prompt level with a range of 0–4 possible points per word. At each session, children could receive a maximum score of 16 points for the production probe (4 points each for 4 words).

Target Vocabulary Test. Vocabulary learning in the Story Friends intervention was measured with a target vocabulary test. Target vocabulary words were those explicitly taught in the intervention. For each word, children were asked to respond to open-ended definitional questions (e.g., *Tell me, what does curious mean?*). When children did not provide a correct response to this item, research assistants provided the standard prompt of *"Curious means…."* All responses were transcribed in real time, audio recorded, and scored at a later time using a three-point scale. A complete definition (e.g., *you want to know more*) or accurate synonym received a score of 2, a partial or associated response received a score of 1 (e.g., *you are curious because you don't know*), and an unrelated or "I don't know" response received a 0. Four words were taught in each book, meaning that the maximum total word points per book was 8 (2 points each for 4 words). The testing schedule for the target vocabulary tests was designed to reduce the amount of testing children completed each week and maximize sensitivity to learning. At the beginning of each unit, children were pretested on all 12 words in the unit. After each week of intervention, children completed the posttest for the 4 words in that week's book. To describe vocabulary learning, gain scores were calculated by subtracting the number of word points at pretest from the number of word points at posttest. The gain score was divided by the total possible word points for that child to provide a percentage of vocabulary learning. Consistent with previous studies of Story Friends, pretest scores were generally very low (average of 8% of possible word points).

In the current study, children completed different numbers of books (between 3–11) meaning that children were taught a different number of target vocabulary words (between 12 and 44). To compare learning in vocabulary intervention across children, a percentage was calculated by the number of word points gained divided by the number of possible word points for each child (2 points per word taught) and multiplied by 100.

General Procedures

All measures were administered by trained undergraduate and graduate research assistants. The CELF-P, PPVT, and EWL were completed prior to intervention, and the target vocabulary measure was completed weekly during intervention. The EWL was given on two sessions on different days to measure word learning proficiency across multiple exposures. All children completed one administration of the EWL; 14 of the 16 children completed a second session. The average length of time between first and second sessions was 2.5 days (range 1–6).

Interrater Reliability

Prior to independent administration, research assistants were trained on each measure by the author and were observed during testing to ensure reliable administration of all measures. Administration of the EWL was audio recorded for the majority of children. To examine fidelity of administration, a trained research assistant listened to 25% of teaching trials and probes using a checklist that captured key elements of administration (i.e., number of presentations of the word, definition). Fidelity of administration was 94% (76–100%). Deviations from administration were typically repetitions of words or phrases, often to maintain children's attention or encourage a response.

All measures were scored by two trained research assistants. For the EWL, no differences were observed between scorers on the definitional and production probes. Given the constraints on possible responses (i.e., little interpretation on the part of the scorer), this high level of agreement is not surprising. For the vocabulary test, a scoring rubric was created for each word that included sample responses that would receive two points, one point, or no points. All responses were transcribed into a spreadsheet so that scorers would be blind to pretest/posttest and scored by a primary and secondary scorer; agreement was 96%. The small number of discrepancies were resolved by a third scorer.

RESULTS

Data Analysis

Dependent variables were standard scores on the PPVT-IV, scores on the definitional and production probes of the EWL at two sessions, and the percentage of words learned in the vocabulary intervention. One child did not complete the PPVT-IV and two children did not complete the second administration of the EWL. These children were included in the analyses. Means and standard deviations for the measures are reported in Table 1. For descriptive purposes, means for children from the Head Start classrooms and the private preschool classroom are reported separately. In this small dataset, there was substantial variability in the dependent variables. Vocabulary knowledge, as measured by the PPVT-IV, ranged from well below to well above average (standard scores of 71–125). On the EWL, children scored at the top and bottom of the possible range for both the definitional and production probes. Learning in vocabulary intervention ranged from 3% to 78% of possible word points.

Table 1 Participant Performance on a Measure of Vocabulary Knowledge, A Process-Based Assessment of Word Learning, and Learning in Vocabulary Intervention

	All Participants ($n = 16$) *M (SD)*	Head Start ($n = 10$) *M (SD)*	Private Preschool ($n = 6$) *M (SD)*	*d*
PPVT-IV	100.25 (18.58)	91.10 (17.22)	115.50 (7.42)	1.84
Definitional 1	3.44 (2.87)	2.70 (2.06)	4.67 (3.78)	.65
Production 1	3.88 (2.90)	3.00 (2.89)	5.17 (1.94)	.88
Definitional 2	5.14 (4.09)	4.75 (3.92)	5.67 (4.63)	.21
Production 2	7.36 (4.27)	5.63 (4.66)	9.67 (2.42)	1.08
Vocabulary Learning	35.69% (20.63)	24.87% (13.69)	51.04% (20.64)	1.49

Note. PPVT-IV: Peabody Picture Vocabulary Test-IV standard score; Definitional Probe 1, 2: Definitional probe of the Explicit Word Learning measure at the first and second sessions, maximum of 12; EWL Production Probe 1, 2: Production probe of the Explicit Word Learning measure at the first and second sessions, maximum of 16, Vocabulary Learning: Percentage of word points gained during intervention.

Preliminary Analysis of EWL

A previous study of the EWL provided preliminary evidence of the validity of the measure for estimating word learning proficiency.[16] Because the EWL is a new measure with only preliminary data, a first step in the current study was to conduct a descriptive analysis to examine group performance on the EWL, identify potential floor or ceiling effects, and determine whether the patterns of performance were similar to the previous study of the measure. Performance on the EWL was examined within and across sessions to determine the information provided by a single administration and to evaluate the information added by a second administration. Scores on the production and definitional probes for Sessions 1 and 2 are presented in Fig. 1. At the first session, there was substantial variation within children with scores ranging between 0–12 on the definitional probe and between 0–9 on the production probe. Similar variation was observed at the second session, the range of scores on the definitional probe was 0–12 and 0–14 on the production probe. For the children who completed two sessions, performance increased from the first to second session for 9 of 14 children on the definitional probe and for 10 of the 14 children on the production probe. Average gains on the definitional probe were 1.86 points (range -4–8 of a possible 12). Performance increased for 10 of the 14 children on the production probe with average gains of 3.36 points (range 0–9). In summary, consistent with the previous study, no floor or ceiling effects were observed and most children increased in performance between the first and second session.

To examine internal consistency of the measure, Cronbach's α was calculated for scores at the first and second sessions. The internal consistency among items was high in both sessions, $\alpha = 0.90$. Test-retest reliability also was high; the correlation between total scores at the first and second sessions was 0.87, $p < .01$.

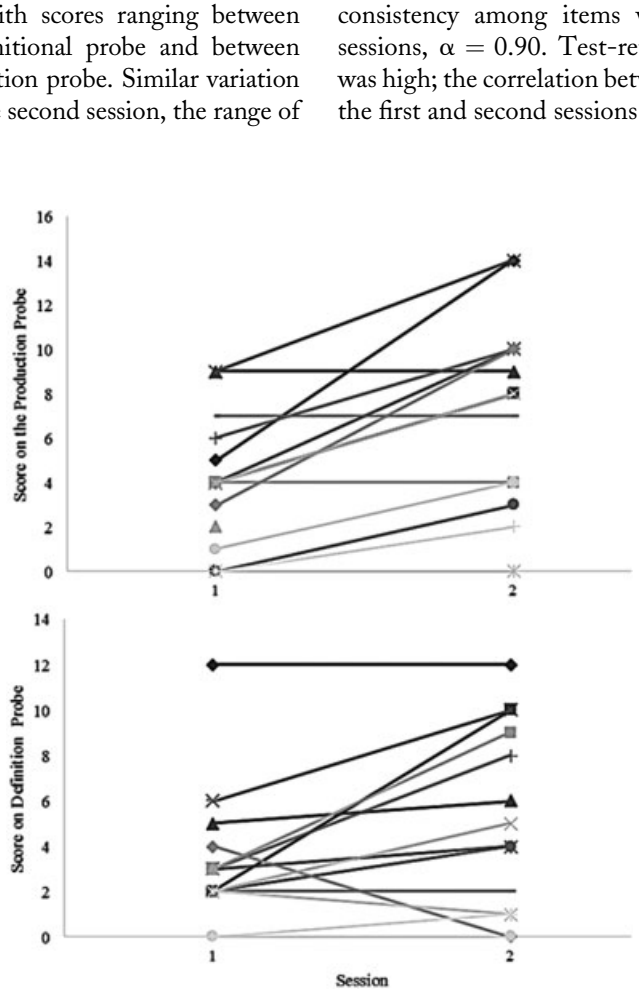

Figure 1 Scores on the production and definitional probes of the explicit word learning measure across two sessions. Each line represents a score for an individual participant.

Relations among Vocabulary Knowledge, Process-based Assessment, and Learning in Vocabulary Intervention

To address the first research question, bivariate correlations were conducted among the PPVT-IV, EWL, and learning in vocabulary intervention (Table 2). Correlations are interpreted here based on Cohen[47] who states that correlations greater than 0.50 are large. Thus, large, positive correlations were found for standard scores on the PPVT-IV, definitional and production probes of the EWL at both sessions, and with learning in vocabulary intervention. As hypothesized, performance on the definitional and production probes of the EWL was significantly and positively correlated with learning in vocabulary intervention, with stronger correlations for performance at the second session than the first session.

Process-Based Assessment of Word Learning for Clinical Decision Making

The next section includes two exploratory analyses that examined the ways in which a SLP might use a process-based assessment of word learning to guide treatment decisions. One purpose of a process-based assessment would be to provide information about word learning proficiency, used in combination with a static measure, to make decisions about intervention intensity. Another purpose would be to provide detailed information about the word learning process to inform intervention approach.

To examine the potential of a process-based assessment to inform intervention intensity, the information provided by scores on the PPVT and EWL were considered. In the current study, six children had standard scores on the PPVT-IV at least one standard deviation below the normative mean, which might be used clinically to identify a child who may need vocabulary instruction. Of these six children, two had very low scores on the first session of the EWL (scores of 0 and 1) and did not improve substantially at the second session (scores of 3 and 4). These low scores indicate that these two children did not learn from the explicit teaching trials on the process-based assessment. Further, these two children were poor responders to the vocabulary intervention, with learning in vocabulary intervention more than one standard deviation below the group mean. The other four children with low PPVT scores had higher scores on the EWL (*M* of 3.75 at the first session and 7.0 at the second session) and demonstrated learning in vocabulary intervention within one standard deviation of the group mean.

The second exploratory analysis examined how the information provided by the process-based assessment could inform decisions about intervention approach. The goal of the EWL was to describe partial knowledge and to be sensitive to learning. Responses on both probes were examined to determine if the incremental scoring and hierarchical prompting were successful in capturing this information. If children scored only at the top and bottom of the range of possible scores (e.g., scores of 0 and 3 on the definitional probe), this would be an indication that the measure only captured incorrect and correct responses. In contrast, if children frequently received scores in the middle of the range

Table 2 Correlations between Measures of Vocabulary Knowledge, Word Learning, and Learning in Response to Intervention

	Definitional 1	Production 1	Definitional 2	Production 2	Vocabulary Learning
PPVT-IV	.53*	.70**	.62*	.67**	.74**
Definitional 1	*	.49	.66*	.74**	.65**
Production 1		*	.57*	.77**	.51*
Definitional 2			*	.60*	.75**
Production 2				*	.66*

Note: **. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 (2-tailed).
Definitional Probe 1: Definitional probe of the Explicit Word Learning measure at the first session; EWL Production Probe 1: Production probe of the Explicit Word Learning measure at the first session, Vocabulary Learning: Percentage of word points gained during intervention.

(e.g., scores of 1 or 2 on the definitional probe, this would be an indication that the EWL was useful to describe partial knowledge. If there were differences in scores between the first and second probes, this would be an indication that the EWL was sensitive to learning.

The incremental scoring on the definitional probe was successful in capturing partial knowledge and was sensitive to learning. On the definitional probe, the maximum score was 3 points if a child responded correctly to the open-ended definitional question and to at least 3 of the 4 yes/no questions. A score of 2 was given if only the open-ended question was correct, and a score of 1 was given if the open-ended question was incorrect and at least 3 of the 4 yes/no questions were correct. At the first probe, scores ranged between 0–3 and most frequently were 0 (40%) or 1 (43%) with only 5% of responses receiving 2 points and 12% receiving 3 points. At the second probe, scores also ranged between 0–3, with an increase in responses that received scores of 2 (13%) and 3 (24%).

Hierarchical prompting on the production probe also captured partial knowledge and was sensitive to learning. At both sessions, responses were scored across the possible range of 0–4 points per word. Responses at each prompting level are presented in Fig. 2. Between the first and second session, more children responded correctly with less prompting. At the first session, most frequent responses (42%) were after the indirect model (e.g., *Another word for happy is yame. How do these boys feel? They feel* …). At the second session, children were more likely to respond after the phonological probe (32%) and the number of open-ended responses increased from just 3% at the first session to 21% at the second session.

In addition to partial knowledge and sensitivity to learning, the EWL was designed to provide information about both definitional knowledge and production abilities, as well as to describe learning of multiple word types. Responses on the EWL were examined to explore these aspects of the measure. As explained in the preliminary analysis section, most

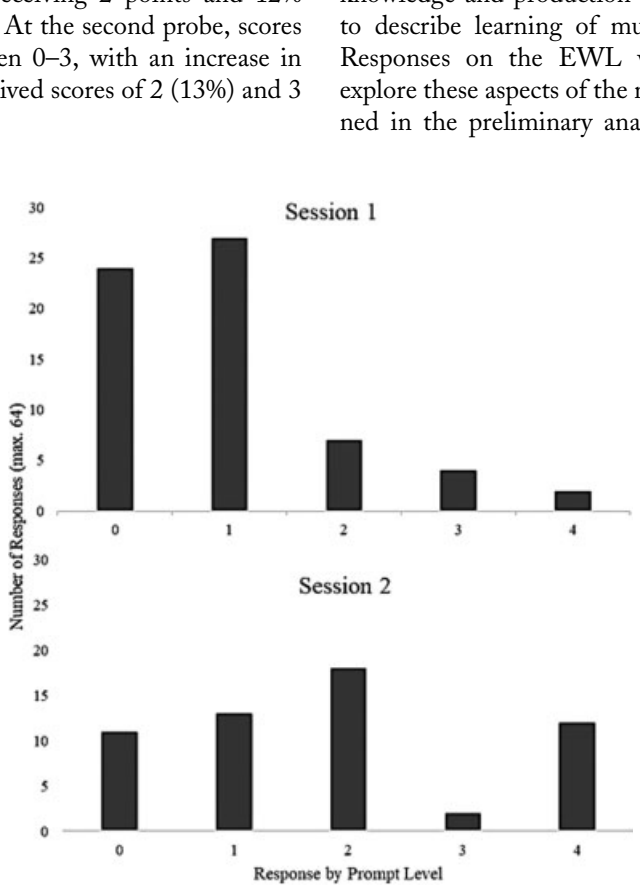

Figure 2 Frequencies of responses at individual prompting levels on the production probe across two sessions. The hierarchical prompts were scored such that the first, more challenging prompt (e.g., open-ended) received the highest score whereas the last, most supportive prompt (e.g., indirect model) received the lowest score of one.

children made gains between the first and second sessions and gains for the production probe were slightly higher than for the definitional probe. However, there was a wide range of performance and individual children had different patterns of responses. For example, one child gained 6 points on the definitional probe but no points on the production probe. Another child demonstrated the opposite pattern, gaining 1 point on the definitional probe and 6 on the production probe. These types of learning patterns might be clinically useful to determine individual strengths and weaknesses. The EWL included both verbs and adjectives; however, no differences in learning of the two-word types was observed.

DISCUSSION

A primary purpose of this study was to examine the relation among vocabulary knowledge, word learning proficiency, and learning in vocabulary intervention. Correlational analyses were conducted among a static, knowledge-based measure of vocabulary knowledge, a process-based assessment of word learning, and learning in vocabulary intervention. Although previous studies have included measures of vocabulary knowledge and word learning or measures of vocabulary knowledge and learning in vocabulary intervention, the current study included a unique dataset with measures of all three components.

As hypothesized, large correlations were observed among vocabulary knowledge, word learning proficiency, and learning in vocabulary intervention. Exploratory analysis indicated that information provided by a process-based measure, when considered in combination with a static measure of vocabulary knowledge, might help to identify children who will benefit from intense intervention. Similarly, other studies have reported that process-based assessment of word learning can predict growth in vocabulary over time.[15] The current study adds to this body of research by highlighting the potential of a process-based assessment of word learning to provide a sophisticated understanding of word learning proficiency of young children and perhaps to predict learning in response to intervention.

A strong correlation was also observed between the static measure of vocabulary knowledge (PPVT-IV) and learning in vocabulary intervention. This finding is consistent with two other studies that have reported a relation between vocabulary knowledge, as measured on static, knowledge-based assessments, and learning in vocabulary intervention in which children with higher vocabulary scores learn more words in intervention.[48,49] An important next step in this line of research will be to examine the extent to which a process-based assessment provides information to predict learning in intervention beyond what is predicted by a static measure.

In this study, children's standard scores on the static measure of vocabulary knowledge were strongly correlated with performance on the process-based assessment of word learning proficiency (EWL). This finding suggests that there was a relation between the words that children knew and the proficiency with which children learned new words. However, in other studies, scores on static, knowledge-based measures of vocabulary and performance on a word learning task have not been related,[19,21,23,50] suggesting that these two types of measures may be assessing different things. In Camilieri and Law,[15] scores on a dynamic assessment were more predictive of vocabulary growth for children in the lower range of vocabulary knowledge than for children in the higher range, leading authors to argue that a dynamic assessment might be particularly useful for children with limited vocabulary knowledge. Across studies, there is evidence to suggest that measures of vocabulary knowledge and process-based assessments of word learning may provide unique information, meaning that including both types of measures may be appropriate. In clinical settings, it is useful to know about both the words the children know and the ways that children approach word learning to make informed decisions about instruction and intervention.

Applications for Clinical Practice

Although the small sample size in the current study prohibits generalization, the findings of two exploratory analyses suggest that a process-based measure of word learning could be useful

to SLPs to make decisions about intervention intensity and approach. First, a process-based assessment might help SLPs match children to appropriately intense interventions. In the current study, of the 6 children who had below-average vocabulary knowledge, two children had the very lowest scores on the EWL at the first session (total scores of 0, 1) and demonstrated very little change at the second session (3, 4). These two children also had the least learning in vocabulary intervention, gaining just 2% and 11% of word points overall. This finding suggests that SLPs who consider information from both a static measure of vocabulary knowledge and a process-based assessment of word learning might identify children who will require intense intervention.

As an example, particularly within groups of children from families with low SES, SLPs might consider a process-based assessment of word learning as a way to distinguish between children who have limited vocabulary as the result of experience and children who have poor word learning proficiency. SES-related differences have been reported for static, knowledge-based measures of single-word vocabulary.[33] These differences have been largely attributed to differences in language-learning experiences.[32] To complement static measures of vocabulary knowledge, process-based assessments can provide an assessment of word learning that may be less dependent on children's previous experience. A child with low vocabulary knowledge but strong word learning proficiency is more likely to benefit from language enrichment to learn words. In contrast, a child with low vocabulary knowledge and poor word learning proficiency might require more intense intervention. Increasing the intensity of intervention may improve learning in vocabulary intervention for poor word learners. This would require more explicit instruction, such as increasing the number of learning opportunities, providing definitions and more contexts for explaining words, and prompting responses.

For children receiving intervention, SLPs might incorporate a process-based assessment into clinical practice to make decisions about intervention approach. The second exploratory analysis indicated that the incremental scoring and hierarchical prompting was useful for describing partial knowledge and captured small changes in knowledge of the words between the two sessions. An SLP might use a similar approach to monitor progress in intervention. Rather than scoring a child's response as correct or incorrect, more sensitive information about small changes in knowledge would help SLPs determine if children were learning in response to intervention.

A process-based assessment can provide information about learning across multiple word types. SLPs might consider using the procedures from the definitional and production probe to measure learning of vocabulary targets from treatment. Because verbs and adjectives are not as easy to depict as nouns, it can be challenging to assess children's learning of these word types; the incremental scoring and hierarchical prompting procedures from the EWL may be a potentially useful clinical tool. SLPs might also compare learning of different word types to determine how to focus intervention. For example, if a child readily acquires easily picturable nouns and action verbs, intervention might focus on more challenging vocabulary such as cognitive state verbs.

SLPs also might compare scores on the definitional and production probes to identify strengths and weaknesses in semantic and phonological knowledge that could inform the intervention approach. For example, a child who demonstrates an ability to define new words but not produce new words might benefit from intervention that strengthened phonological representations of words, perhaps through increasing exposures to the word. In contrast, a child who readily produces new words but struggles to define them might be better suited to an intervention that focuses on semantic knowledge including associations, synonyms and antonyms, and categories.

Limitations

One important limitation of the current study was the small sample size. Although this dataset was appropriate for the purposes of this paper, the sample size prevented the use of more sophisticated analyses. In a larger sample, a regression analysis might better explain the relative contributions of measures of vocabulary

knowledge and process-based assessments to the prediction of learning in vocabulary intervention. Another limitation was that limited demographic information was available to describe participants. Although center enrollment provided a rough indication of socioeconomic status, more detailed information would have been useful to describe individual children.

The data in the current study were collected as part of a pilot study evaluating a new component of the *Story Friends* intervention. Although other studies of *Story Friends* have indicated that the treatment is generally effective in improving vocabulary knowledge in preschool children, the books and words included in this pilot study have not yet been subject to rigorous evaluation. Because children varied in the number of weeks they participated in the intervention, the dependent variable for learning in vocabulary intervention was the percentage of word points gained of the possible total word points for each child. However, it is not possible to determine whether each week of intervention was equivalent in difficulty (e.g., some weeks may have included more difficult words), meaning that the comparison between children can only be an estimate.

CONCLUSIONS

Findings of this preliminary study suggest that vocabulary knowledge, word learning proficiency, and learning in vocabulary intervention are related. A process-based assessment of word learning may help indicate which children will demonstrate learning in vocabulary intervention. Exploratory analyses contribute to a larger body of work that highlights the potential contributions of a process-based assessment of word learning to clinical decision making by SLPs.

ACKNOWLEDGMENTS

This research was partially supported by grants from the United States Department of Education, Institute of Education Sciences, R432A150132, awarded to the University of South Florida, Howard Goldstein and Elizabeth Kelley, PIs and R324C08001 awarded to the University of Kansas, Charles Greenwood and Judy Carta, PIs.

DISCLOSURE

Elizabeth Kelley and Howard Goldstein are authors of Story Friends and have a financial interest, as they receive royalties from sales of this product through Paul Brookes Publishing.

REFERENCES

1. National Early Literacy Panel. Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy; 2008
2. Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: evidence from a longitudinal structural model. Dev Psychol 2002; 38(06):934–947
3. Walker D, Greenwood C, Hart B, Carta J. Prediction of school outcomes based on early language production and socioeconomic factors. Child Dev 1994;65(2 Spec No):606–621
4. Dickinson DK, Porche MV. Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. Child Dev 2011;82(03): 870–886
5. Dunn LM, Dunn DM. Peabody Picture Vocabulary Test - IV. 4th ed. Minneapolis, MN: NCS Pearson Assessments; 2007
6. Gutie Rrez-Clellen VF, Pen A E. Dynamic assessment of diverse children: A tutorial. Lang Speech Hear Serv Sch 2001;32(04):212–224
7. Burton VJ, Watkins RV. Measuring word learning: dynamic versus static assessment of kindergarten vocabulary. J Commun Disord 2007;40(05): 335–356
8. Pena E, Iglesias A, Lidz C. Reducing test bias through dynamic assessment of children's word learning ability. Am J Speech Lang Pathol 2001; 10:138–154
9. Ukrainetz TA, Harpell S, Walsh C, Coyle C. A preliminary investigation of dynamic assessment with Native American kindergartners. Lang Speech Hear Serv Sch 2000;31(02):142–154
10. Hasson N, Joffe V. The case for dynamic assessment in speech and language therapy. Child Lang Teach Ther 2007;23:9–25
11. Peña ED, Gillam RB, Bedore LM. Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. J Speech Lang Hear Res 2014;57(06): 2208–2220
12. Kapantzoglou M, Restrepo MA, Thompson MS. Dynamic assessment of word learning skills: identifying language impairment in bilingual children. Lang Speech Hear Serv Sch 2012;43(01):81–96

13. Camilleri B, Botting N. Beyond static assessment of children's receptive vocabulary: the dynamic assessment of word learning (DAWL). Int J Lang Commun Disord 2013;48(05):565–581
14. Camilleri B, Law J. Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. Int J Speech-Language Pathol 2007;9:312–322
15. Camilleri B, Law J. Dynamic assessment of word learning skills of pre-school children with primary language impairment. Int J Speech-Language Pathol 2014;16(05):507–516
16. Kelley ES. Measuring Explicit Word Learning of Preschool Children: A Development Study. Am J Speech Lang Pathol 2017;26(03):961–971
17. Loftus S, Coyne M, McCoach B, Zipoli R. Effects of a supplemental vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties. Learn Disabil Res Pract 2010;25:124–136
18. Pullen PC, Tuckwiller ED, Konold TR, Maynard KL, Coyne MD. A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disability. Learn Disabil Res Pract 2010; 25:110–123
19. Rice ML, Oetting JB, Marquis J, Bode J, Pae S. Frequency of input effects on word comprehension of children with specific language impairment. J Speech Hear Res 1994;37(01):106–122
20. Storkel HL, Voelmle K, Fierro V, Flake K, Fleming KK, Romine RS. Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: Identifying an adequate intensity and variation in treatment response. Lang Speech Hear Serv Sch 2017;48(01): 16–30
21. Kiernan B, Gray S. Word learning in a supported-learning context by preschool children with specific language impairment. J Speech Lang Hear Res 1998;41(01):161–171
22. Gray S. Word learning by preschoolers with specific language impairment: predictors and poor learners. J Speech Lang Hear Res 2004;47(05): 1117–1132
23. Gray S. The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. J Speech Lang Hear Res 2006;49(05): 955–969
24. Dollaghan CA. Fast mapping in normal and language-impaired children. J Speech Hear Disord 1987;52(03):218–222
25. Ellis Weismer S, Hesketh LJ. Lexical learning by children with specific language impairment: effects of linguistic input presented at varying speaking rates. J Speech Hear Res 1996;39(01):177–190
26. Ellis Weismer S, Hesketh LJ. The impact of emphatic stress on novel word learning by children with specific language impairment. J Speech Lang Hear Res 1998;41(06):1444–1458
27. Gray S. Word learning by preschoolers with specific language impairment: effect of phonological or semantic cues. J Speech Lang Hear Res 2005;48 (06):1452–1467
28. Alt M, Plante E. Factors that influence lexical and semantic fast mapping of young children with specific language impairment. J Speech Lang Hear Res 2006;49(05):941–954
29. Alt M, Plante E, Creusere M. Semantic features in fast-mapping: performance of preschoolers with specific language impairment versus preschoolers with normal language. J Speech Lang Hear Res 2004;47(02):407–420
30. Gray S, Brinkley S. Fast mapping and word learning by preschoolers with specific language impairment in a supported learning context: effect of encoding cues, phonotactic probability, and object familiarity. J Speech Lang Hear Res 2011;54(03):870–884
31. Gray S. Word-learning by preschoolers with specific language impairment: what predicts success? J Speech Lang Hear Res 2003;46(01):56–67
32. Hoff E. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. Child Dev 2003;74 (05):1368–1378
33. Qi CH, Kaiser AP, Milan S, Hancock T. Language performance of low-income African American and European American preschool children on the PPVT-III. Lang Speech Hear Serv Sch 2006;37 (01):5–16
34. Washington JA, Craig HK. Performances of At-Risk, African American Preschoolers on the Peabody Picture Vocabulary Test-III. Lang Speech Hear Serv Sch 1999;30(01):75–82
35. Horton-Ikard R, Ellis Weismer S. A preliminary examination of vocabulary and word learning in African American toddlers from middle and low socioeconomic status homes. Am J Speech Lang Pathol 2007;16(04):381–392
36. Olswang LB, Bain BA. Assessment information for predicting upcoming change in language production. J Speech Hear Res 1996;39(02):414–423
37. Wiig EH, Secord WA, Semel E. Clinical Evaluation of Language Fundamentals Preschool - 2. 2nd ed. San Antonio, TX: Harcourt Assessment; 2004
38. Kelley ES, Goldstein H, Spencer TD, Sherman A. Effects of an automated tier 2 storybook intervention on vocabulary and comprehension learning in preschool children with limited oral language skills. Early Child Res Q 2015;31:47–61
39. Goldstein H, Kelley E, Greenwood C, Carta J, Atwater J, McCune L. Embedded Instruction Improves Vocabulary Learning during Storybook

Reading among High-Risk Preschoolers. Manuscript in preparation 2015
40. Spencer E, Goldstein H, Sherman A, et al. Effects of an automated vocabulary and comprehension intervention: An early efficacy study. J Early Interv 2012;34:195–221
41. Greenwood C, Carta J, Kelley E, et al. The effects of a Tier 2 vocabulary and comprehension storybook intervention on preschool children's early learning: A replication. Elem Sch J 2016;116:574–599
42. Kelley ES, Goldstein H. Building a Tier 2 intervention: A glimpse behind the data. J Early Interv 2015;36:292–312
43. Goldstein H, Kelley E, Greenwood C, et al. Embedded instruction improves vocabulary learning during automated storybook reading among high-risk preschoolers. J Speech Lang Hear Res 2016;59(03): 484–500
44. Storkel HL, Armbrüster J, Hogan TP. Differentiating phonotactic probability and neighborhood density in adult word learning. J Speech Lang Hear Res 2006;49(06):1175–1192
45. Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. Behav Res Methods 2012;44(04): 978–990
46. Beck I, McKeown M. Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. Elem Sch J 2007;107:251–272
47. Cohen J. A power primer. Psychol Bull 1992;112 (01):155–159
48. Marulis LM, Neuman SB. The effects of vocabulary intervention on young children's word learning: A meta-analysis. Rev Educ Res 2010; 80:300–335
49. Justice LM, Meier J, Walpole S. Learning new words from storybooks: an efficacy study with at-risk kindergartners. Lang Speech Hear Serv Sch 2005;36(01):17–32
50. Rice ML, Buhr J, Oetting JB. Specific-language-impaired children's quick incidental learning of words: the effect of a pause. J Speech Hear Res 1992;35(05):1040–1048