“Show Me X”
New Views of an Old Assessment Technique

Mabel L. Rice and Ruth V. Watkins

The mother of a 3-year-old girl holds her child in her lap as they look at a picture book together. The mother says to her daughter, “Show me the kitty. Where is the kitty?” When the child points to the picture of the kitty, the mother says, “That’s right! You know the kitty, don’t you?”

A psychologist wants to determine how intelligent children are. Knowing that an important kind of intelligence is the ability to learn new words and that this ability is evident in young children, the psychologist devises a test that consists of showing young children a set of pictures of everyday things. The child is instructed, for example, to “show me kitty.” The number of correct responses shows how many words the child knows.

A behavioral scientist studies the ways in which young children come to know the meanings of words like “kitty.” The scientist introduces 3-year-old children to novel things with novel names, like “blick.” Then the scientist shows a child a set of three pictures—one is a “blick,” and two are other novel things that differ from “blick” in carefully specified ways. The scientist says to the child, “Show me blick,” and carefully records which picture the child chooses. The child’s choice of pictures indicates the ways in which novel things can be interpreted and which novel things go together.

A speech-language pathologist evaluates a 3-year-old girl who does not yet talk very well. It is important to know how much the girl understands of the language she hears. The speech-language pathologist uses a test consisting of a book with four pictures per page and asks the girl, for example, to “show me kangaroo.” The pathologist then calculates a test score that will determine whether the child’s performance is within the range expected for children her age.

What these scenarios illustrate is that parents, psychologists, behavioral scientists, and speech-language pathologists all rely on the same simple proce-
dare for assessing young children’s understanding of words. Binet and Simon, in 1916, advocated this procedure with their now-famous intelligence test:

One of the best tests upon which to form a judgment [about the intelligence of individuals] is to ask them to designate in a picture the object which one names for them. The test is so much the more to be recommended because it has, for the normal child, the great attraction of curiosity. There is also a great advantage in asking him to point out the objects corresponding to the words which are said to him, rather than to make him name the objects which he himself sees, because of the defects of pronunciation which often prevent him from being understood. (p. 145)

Although “show me X” has long been considered a useful technique for assessing children’s word comprehension, the contemporary literature introduces innovative ways of understanding lexical acquisition and language impairments. This chapter examines contemporary developments relevant to the third and fourth scenarios previously described; that is, using “show me X” to learn about the mechanisms and processes of lexical acquisition and to assess children with language impairments. In some ways, the available research is not fully in accordance with the conventional wisdom about this measure. In short, “show me X” has the advantage of ease and familiarity, but it also is a more subtle and complex assessment means than is self-evident.

BACKGROUND

“Show me X” methods for assessing vocabulary comprehension abilities have a long history, dating to at least 1900 (cf. Binet & Simon, 1916; Terman, 1916). Such vocabulary comprehension methods continue to be a component of several measures of intelligence (e.g., Stanford-Binet Intelligence Scale [Thurstone, Hagen, & Sattler, 1986], Kaufman Assessment Battery for Children [Kaufman & Kaufman, 1983]). Furthermore, the picture-identification method of evaluating vocabulary comprehension is a staple of conventional clinical assessment tools. The Peabody Picture Vocabulary Test–Revised (PPVT–R) (Dunn & Dunn, 1981) is the prototype vocabulary comprehension measure; the same basic task is used to assess vocabulary comprehension in a variety of other instruments, including the Receptive One-Word Picture Vocabulary Test (Gardner, 1985), as well as subtests of the Test of Language Development–P:2 (Newcomer & Hammill, 1991) and the Test for Auditory Comprehension of Language–Revised (Carrow-Woolfolk, 1985).

The motivations behind the widespread use of picture-identification methods of evaluating vocabulary comprehension are relatively straightforward. First, assessing vocabulary comprehension addresses the importance of lexical development. Children’s comprehension of words is fundamental to children’s emerging linguistic system. Understanding the meaning of words is a hallmark indicator of the beginnings of language. At very young ages, the ways in which children quickly learn new word meanings reveal much about the nature of their early lexical development and related cognitive abilities. Beyond the first stages, children’s word comprehension serves as an indicator of lexical achievement that is interpretable as one important aspect of general intelligence and as a contributor to children’s literacy.

Second, there are methodological advantages to the “show me X” method:

1. As noted by Binet and Simon (1916), the procedure capitalizes on children’s inherent interest in looking and pointing; in short, children like to do this activity.

2. Young children often cannot or will not name objects; therefore, the pointing response is easier and probably more accurately elicited, especially for young children. It also is likely that comprehension of a particular word can be evident before a child may be able to say the word, presumably because the necessary phonetic representations for pronunciation may require greater experience with the words (cf. Bates, 1993).

3. Although the picture-identification procedure is suitable for and of interest to very young children, it also is appropriate for older children and adults. This makes it one of the few measures of language ability that can be employed across the life span.

4. Manipulating the foils can reveal much about the basis for a child’s judgments about word meanings. Many studies in the contemporary literature have capitalized on this feature to increase what is known about early lexical development.

5. Such tasks can be used to identify children with language impairments, to study their lexical acquisition processes, to determine their relative growth in vocabulary as compared to their age-peers, and to plan for intervention.

This chapter addresses four specific issues: 1) the role of vocabulary comprehension tasks in revealing powerful processes and mechanisms of lexical acquisition; 2) using vocabulary comprehension tasks to detect lexical impairments and to reveal mechanisms of lexical development in children with specific language impairment (SLI); 3) using vocabulary comprehension as a predictor of developmental outcome for children with language impairments, as well as applying lexical comprehension measures to intervention efficacy studies; and 4) interpretive boundaries: that is, some surprising dissociations between vocabulary comprehension abilities and other linguistic skills, as well as limitations in the applicability of particular measures of lexical knowledge. The chapter concludes with a summary of issues in current and future applications of vocabulary comprehension measures in research and clinical contexts.

The main points are these: Just as hammers and screwdrivers are valuable tools in a carpenter’s tool kit, so the “show me X” way of assessing vocabulary comprehension is a valuable tool in the armamentarium of researchers and practitioners. It is worthwhile, from time to time, to step back and reconsider the full range of usefulness of such procedures. The literature review and synthesis pro-
vided here is intended to yield that reconsideration. Just as hammers and screwdrivers cannot be applied indiscriminately to any woodworking problem, the “show me X” procedure is not without limitations. This chapter identifies some of those limitations and offers caveats where appropriate.

INSIGHTS FROM VOCABULARY COMPREHENSION TASKS ON LANGUAGE ACQUISITION AND IMPAIRMENT

Insights on Language Acquisition

A review of the available research on children’s lexical acquisition reveals that vocabulary comprehension tasks play a central role in the study of processes of children’s lexical learning. The use of vocabulary comprehension tasks has revealed much about young children’s word-learning mechanisms. A large number of investigations have employed both picture-pointing and object-identification methods of measuring comprehension abilities, accompanied by a variety of verbal prompts such as “Where’s X,” “Find X,” “Show me X,” or “Point to X” (Carey & Bartlett, 1978; Dollaghan, 1985, 1987; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Rice & Woodsmall, 1988). These studies have identified several powerful lexical acquisition capabilities including fast mapping, quick incidental learning, and mechanisms of word comprehension.

Fast Mapping  Investigations of children’s lexical comprehension demonstrate that young children can “fast map” new word meanings, forming initial (albeit somewhat sketchy) representations of a new word’s meaning. This allows children to quickly build a lexicon of thousands of words by the time they enter school (Carey, 1982). Carey and Bartlett (1978; see also Carey, 1978, 1982) were among the first to document young children’s impressive word-learning aptitudes. Carey and Bartlett used a relatively naturalistic vocabulary comprehension task to probe children’s ability to learn new words. In this task, an adult presented children with two trays, one of a known color (e.g., red or blue) and the other of an unknown color (e.g., olive green). Children were then told, “Find the chromium tray. Not the red one, the chromium one.” The phenomenon observed by Carey and Bartlett is widely recognized. Preschoolers were adept in fast mapping initial properties of a word’s meaning from just one or two exposures to that word in a meaningful context; that is, young children were able to select the chromium tray without any explicit teaching of the novel word. A number of subsequent investigations confirmed children’s aptitude for rapidly mapping new word meanings from limited experience with those words (Dickinson, 1984; Dollaghan, 1985; Heibeck & Markman, 1987; Rice & Woodsmall, 1988).

Quick Incidental Learning  Vocabulary comprehension tasks have served as the vehicle for additional insights into children’s word-learning mechanisms. For example, Rice and Woodsmall (1988) evaluated 3- and 5-year-old children’s ability to map novel word meanings in a noninteractive context—video viewing. In this task, children were exposed to new object, action, attribute, and affective state words embedded in narration specifically developed to match a cartoon segment. Each novel word was used approximately five times during the course of the video cartoon. Children’s comprehension of target words was assessed, pre- and post-video viewing, through a vocabulary comprehension task analogous to the PPVT–R; pictures taken from the video program were used in an array of four photographs with a “show me X” prompt. Rice and Woodsmall found that both 3- and 5-year-olds learned new words through video viewing, with greater learning demonstrated by the 5-year-old children.

The context for introducing novel words and evaluating children’s learning was interactive in studies that preceded Rice and Woodsmall (1988) (cf. Carey & Bartlett, 1978; Dickinson, 1984; Dollaghan, 1985, 1987; Heibeck & Markman, 1987). In those investigations, novel words were presented in focused and presumably salient circumstances; an adult used the target word in a play-based activity, and the child was asked to demonstrate comprehension or knowledge of the new word in a similar manner using real objects (see Dollaghan, 1985, for example). The context for word exposure and the comprehension measure used by Rice and Woodsmall differed in fundamental ways from the majority of fast-mapping investigations. The Rice and Woodsmall study revealed that young children’s word-learning mechanisms are robust enough to accomplish at least partial mappings of new word meanings through strictly incidental exposure to novel forms. Hence, Rice (1990, 1991) suggested, a quick incidental learning (QUIL) capability exists in children’s word-acquisition competencies. This finding is important because it reveals that children can learn new words in indirect experiential contexts, such as television viewing, in addition to directly focused interactions with an adult. Television’s influence on word learning is well-documented (cf. Rice, Huston, Truglio, & Wright, 1990).

Mechanisms of Word Comprehension  Evidence of children’s powerful word-learning aptitudes and accomplishments has led scholars to consider mechanisms that may underlie and promote lexical achievements (Golinkoff et al., 1994). Specifically, investigators have proposed and evaluated particular principles or constraints thought to enable rapid word acquisition. The basic proposal is that these principles operate to reduce the amount of information that must be considered in determining a word’s meaning (Baldwin & Markman, 1989; Clark, 1990; Golinkoff, Hirsh-Pasek, Bailey, & Wexner, 1992; Golinkoff et al., 1994; Markman & Wachtel, 1988; Merriman & Bowman, 1989; Mervis & Bertrand, 1993; Mervis, Golinkoff, & Bertrand, 1994; Waxman & Kosowski, 1990). For example, Golinkoff et al. (1994) summarized fundamental principles that direct children’s lexical acquisition, principles such as reference (i.e., words are mapped onto representations of objects); extendibility (i.e., a word can refer to objects beyond a labeled exemplar); and object scope, or whole object (i.e., the first referent considered for a word is an entire object...
rather than its parts or attributes). These basic lexical principles are believed to be accessible to the young child at about 1 year of age. The principles focus the learner, constrain the hypotheses that the learner will entertain in attempting to infer the meanings of novel words, and, thus, narrow the scope and magnitude of the word-learning task.

A number of specific lexical principles have been proposed, some offering contradictory perspectives. Investigators continue to actively debate the status of such principles, accumulating evidence and negotiating the particulars of the various principles (cf. Markman, 1989; Mervis et al., 1994). One aspect of this debate, however, is abundantly clear: Vocabulary comprehension tasks, using either objects or pictures, have played a crucial role in understanding and evaluating proposed lexical principles. Research assessing two particular principles illustrates this point. Markman and her colleagues (Baldwin & Markman, 1989; Markman, 1989; Markman & Wachtel, 1988) proposed that a mutual exclusivity (ME) bias exists in young word learners, suggesting that children assume that an object will have only a single name. On this account, a novel label is attached to an unnamed object, and the children refrain from attaching a second label to an already named object. In situations in which a novel name occurs, and the surrounding context suggests that the novel name refers to a previously named object, Markman and her colleagues anticipated that ME will take precedence over the whole object principle described previously; that is, the child will assume that the novel name refers to a part or attribute of the object.

In contrast, Golinkoff et al. (1994; Mervis & Bertrand, 1993, 1994) suggested a lexical principle termed novel name-nameless category (N,C). N,C states simply that young children are able to associate novel terms and previously unnamed objects rapidly, presumably after one or two exposures in incidental or nonexplicit contexts. The onset of the principle is posited to be about 18 months—around the time the child begins to learn words at a brisk pace. Prior to that time, word acquisition typically proceeds at a relatively slow rate (cf. Golinkoff et al., 1994).

These components of N,C are not in conflict with Markman’s ME principle. In fact, in contexts in which a child encounters a novel word and an unnamed object, N,C and ME both anticipate that the child will map the new term to the whole unnamed object. The two perspectives differ in situations in which the child hears a novel word in the presence of an already named object. Whereas ME suggests that the child will attach the new label to a salient attribute or component of the already named object, N,C makes no specific predictions about likely word-referent pairings in a situation in which a named object receives a second label. Instead, Golinkoff et al. (1994) suggested that the child will evaluate candidate interpretations of the new term in accordance with available linguistic and nonlinguistic contextual information. For example, only if a child heard the novel word while viewing the manipulation of a part of an already named object would the child be likely to infer that the label referred to the object part (cf. Baldwin, 1993). Other plausible referents would be equally likely, as directed by salient contextual information. For example, if the object were a particularly unusual or previously unidentified color, the novel word might be hypothesized to be a color label. If reasonable candidate referents were not readily apparent, then the child would probably accept a second label for an already named object. Thus, a primary difference between the two lexical acquisition principles is that ME asserts that the child will automatically apply a competing label to a part or component of an already named object, whereas N,C allows more degrees of freedom and contends that the child will consider a range of options, including a second label for an already named object. This learning is guided by the contextual information available as the child chooses a referent for the competing label.

Data obtained through the vocabulary comprehension paradigm (i.e., “show me X”) have helped tease out the relative merits of the ME and N,C perspectives. For example, Markman and Wachtel (1988) tested the predictions of the ME account by assessing children’s willingness to apply a novel name when they heard the new term paired with a known or identified object. In about half of the comprehension trials, Markman and Wachtel found that 3-year-olds selected object components or attributes, and about half of the time they selected the object itself. Because the children were willing to accept a second label in many of the trials used by Markman and Wachtel, Golinkoff et al. (1994) interpreted these findings as inconsistent with the ME principle.

Additional data support the predictions of N,C versus ME. Specifically, Mervis et al. (1994) conducted an experiment similar to Markman and Wachtel’s, again evaluating young children’s comprehension of novel words when the terms were used in the presence of an already named object. Following exposure to novel terms, the children were willing to apply both known and novel words to the same object (i.e., children who knew the word “truck” were willing and able to select a truck when both “show me truck” and “show me lorry” were probed in comprehension). These findings conflict with the expectations of the ME lexical principle.

In summary, accounts of early word acquisition have underscored the role of lexical principles in simplifying the child’s task and explaining rapid learning curves. Key insights into these impressive word-acquisition mechanisms have been obtained through measures of language comprehension. Vocabulary comprehension tasks have uncovered the boundaries of initial word mapping, the process of QUIL, and the existence of lexical principles that expedite word acquisition by focusing children on particular candidate referents for novel word meanings.

Insights on Language Impairment

It is well-documented that children with SLI demonstrate difficulties in lexical acquisition (Leonard, 1988; Rice, 1991). The typical profile of a child with SLI is
a pattern of late appearance of first words, followed by slower-than-expected lexical acquisition through the preschool and early school years (Crystal, 1987). Parallel impairments in vocabulary comprehension also may be a frequent component of the SLI profile (Johnston, 1988; Leonard, 1988; Rice, 1991).

Because many young children with language-learning difficulties evidence vocabulary comprehension impairments, researchers and clinicians use vocabulary comprehension measures as a diagnostic marker of the impairment. Many scholars in the field use below-average PPVT–R standard scores as a criterion in SLI identification and/or rely on the PPVT–R as a key index for subject description (e.g., Craig & Evans, 1993; Craig & Washington, 1993; Kelly & Rice, 1994; Masterson, 1993; Oetting & Rice, 1993; Rice, Buhr, & Nemeth, 1990; Rice, Buhr, & Oetting, 1992; Rice, Oetting, Marquis, Bode, & Pae, 1994; Rice, Wester, & Cleave, 1995; Tomblin, Freese, & Records, 1992; Watkins, Buhr, & Davis, 1993; Watkins, Kelly, Harbers, & Hollis, 1995; Watkins & Rice, 1991; Weismer, 1991; Weismer & Hesketh, 1993).

Research addressing word learning in children with SLI has revealed not only delays in lexical acquisition but also apparent differences in the mechanisms driving development. Rice and her colleagues (Rice, Buhr, et al., 1990; Rice et al., 1992; Rice et al., 1994) have extended the QUIL paradigm to the SLI population as a means to understand processes of impaired lexical development. Again, vocabulary comprehension tasks have played a central role in revealing breakdown in word-learning mechanisms. Rice, Buhr, et al. (1990) replicated the videotape task developed in Rice and Woodsmal (1988) with 5-year-old children with SLI. Results revealed that children with SLI learned fewer words, as measured in a vocabulary comprehension format, than did either language- or age-equivalent peers with typical language development. Children with SLI did demonstrate some QUIL abilities but at a reduced level relative to typical counterparts (i.e., the children with SLI comprehended an average of 1.5 new words, of 20 possible, after the video-viewing activity; language-equivalent peers learned 2.3 new words, and age-equivalent peers learned 4.2 new forms). Rice, Buhr, et al. evaluated plausible sources of the limited ability of children with SLI to quickly comprehend new words; candidate explanations involved restrictions in processing and/or segmenting abilities whereby children with SLI may be unable to identify novel words in the incoming stream of verbalization. This limitation, in turn, could be tied to particular syntactic difficulties, such as impairments in grammatical morphology.

In subsequent investigations, Rice and her colleagues systematically evaluated potential sources of the restricted QUIL aptitudes of children with SLI. These studies extended the use of the vocabulary comprehension format developed by Rice and Woodsmal (1988). Rice et al. (1992) increased the salience of target words in a video-viewing QUIL task by presenting the target words in sentence-final position and adding a brief pause immediately before presentation of the target word. The general aim of this project was to evaluate whether manipulations of target-word salience would enhance the QUIL abilities of children with SLI. Results revealed no advantage for children with SLI or for their typically developing counterparts in the enhanced salience condition. In general, these findings suggest that the restricted QUIL capabilities of children with SLI are not likely to be tied to an inability to segment novel words within the context of the language stream.

Next, Rice et al. (1994) assessed the influence of varied input frequencies and word types on the QUIL abilities of children with SLI and age- and language-equivalent counterparts. Both initial comprehension and later retention of new words were examined. Several intriguing findings were revealed. First, given more instances of target words, children with SLI demonstrated robust QUIL abilities; with 10 exposures to target words, children with SLI showed posttest comprehension gains that were comparable to their age-equivalent peers. Second, contrasts of nouns and verbs in QUIL revealed a surprising pattern of verb learning and loss in children with SLI. During posttesting in the high input frequency condition (10 exposures), children with SLI performed as well on comprehension of novel verbs as did their age-equivalent peers; however, the children with SLI subsequently displayed a sharp loss of verb forms in the retention condition (verb comprehension measured 1 week later). Thus, Rice et al. suggested that increased frequency is necessary to promote lexical comprehension for children with SLI. At the same time, increased input alone is not likely to be sufficient to ensure retention of new word forms. Rice et al. hypothesized that verb learning, specifically long-term storage of verbs in semantic networks, may be influenced by the grammatical markings associated with verb forms (e.g., the inflections associated with them and the grammatical frames in which they appear). In this study, available grammatical markings may not have been particularly helpful in mapping new words. Target verbs were modeled in the video presentation in limited grammatical contexts that consisted predominantly of uninflected forms and forms with past-tense markings. Because the children with SLI and the language-equivalent peers who participated in this study lacked mastery of the regular past-tense inflection, the verb presentation may not have been informative enough to ensure adequate storage for long-term retention.

Finally, Rice, Cleave, Oetting, and Pae (1993) evaluated young children’s ability to use grammatical devices that cue differences in novel mass versus count noun forms. In a QUIL word-learning task, presented via videotape and evaluated in a picture-pointing comprehension format, children with SLI and age- and language-equivalent peers were required to differentiate between novel mass and count nouns, as specified by different determinant cues presented with novel objects and substances (e.g., “a keelwig,” “some blick”). Findings revealed that children with SLI were less capable than were their age-equivalent counterparts in distinguishing mass versus count nouns, making errors in the association between novel label and novel substance or object. In contrast, typically developing 5-year-olds (age-equivalent peers) made use of grammatical
cues in achieving links between novel labels and substances or objects. These findings suggest an interaction between morphosyntactic difficulties and lexical development in children with SLI, as learning new words is constrained by challenges in recognizing grammatical cues to particular word classes.

In summary, measures of vocabulary comprehension are fundamental to the study of child language disorders, serving as a diagnostic indicator of SLI. The PPVT-R is the measure most commonly used for this purpose. Furthermore, within the context of QUIL investigations, measures of vocabulary comprehension that emulate the PPVT-R format have revealed that 1) children with SLI are less able than their peers to map new word meanings, at least in contexts in which incidental, nonsalient exposure to novel forms is provided; 2) such impairments are not likely to be attributable to problems with segmenting language input; 3) with enhancement of input frequency, children with SLI demonstrate QUIL capabilities roughly commensurate with their peers, but they may experience additional difficulties in long-term retention of new forms; and 4) children with SLI are less facile in using grammatical devices to cue particular distinctions in novel word form class when presented in a QUIL context. In brief, these investigations and the vocabulary comprehension format central to them have revealed much about the nature and character of SLI, including insight into the mechanisms of acquisition operative for children with language limitations.

One caution about using the PPVT-R as a diagnostic marker of SLI warrants mention. A high proportion of the items on the PPVT-R are nouns. Given that children with SLI display unique problems in retaining verb forms (cf. Rice et al., 1994; see also Rice & Bode, 1993; Watkins, Rice, & Moltz, 1993), the PPVT-R may, in fact, underestimate the lexical difficulties of children with SLI.

**VOCABULARY COMPREHENSION: DEVELOPMENTAL OUTCOMES, RELATED DOMAINS, AND INTERVENTION STUDIES**

The use of vocabulary comprehension measures has not been limited to the study of typical and atypical lexical development. Vocabulary comprehension measures have been utilized in a considerable number of investigations with a range of aims (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Butler, Marsh, Sheppard, & Sheppard, 1985; Chaney, 1994; Crain-Thoreson & Dale, 1992; Dunning, Mason, & Stewart, 1994; Karweit, 1989; Lonigan, 1994; Rice, Hutton, et al., 1990; Scarborough & Dobrich, 1994; Schleicker, White, & Jacobs, 1991; Tunmer, Herriman, & Nesdale, 1988; Walker, Greenwood, Hart, & Carta, 1994; Whitehurst et al., 1988; Whitehurst et al., 1994). Key applications of vocabulary comprehension measures include 1) predicting developmental trajectories (Bishop & Edmundson, 1987; Rice & Hadley, 1995; Thal & Tobias, 1992; Thal, Tobias, & Morrison, 1991); 2) assessing associations between language skills and other developmental areas, such as social competence, literacy experiences, reading development, and academic achievement (Anderson & Freebody, 1981; Chaney, 1994; Gertner, Rice, & Hadley, 1994; Walker et al., 1994); and 3) evaluating the outcomes of specific and general intervention programs (Arnold et al., 1994; Brody, Stoneman, & McCoy, 1994; Karweit, 1989; Rice & Hadley, 1995; Whitehurst et al., 1988; Whitehurst et al., 1994). In these various investigations, vocabulary comprehension measures have served a number of functions, most often acting as either a variable predicting status in other domains or as an index of developmental outcome.

The widespread use of vocabulary comprehension measures can undoubtedly be attributed to many of the reasons cited at the beginning of the chapter (e.g., they can be used with young children and across the life span, they place minimal demand on the individual tested). The PPVT-R in particular has two additional attributes that make it an appealing choice in large-scale research endeavors that employ multiple measures: It is easy to administer, and it is brief (i.e., the PPVT-R does not require extensive training or practice and typically is completed in 10–15 minutes). Beyond these functional benefits, however, measures of vocabulary comprehension appear to be powerful indices of language ability that have predictive and concurrent associations with multiple other dimensions of developmental competence. Key findings related to three specific applications of vocabulary comprehension measures are highlighted: 1) prediction of continued language impairment and intervention outcomes, 2) concurrent relations between vocabulary comprehension and social competence, and 3) use in intervention studies.

**Developmental Outcomes**

Investigations of early language development have revealed that a substantial number of children demonstrate delayed onset of lexical acquisition and slow word-learning patterns. These children have been identified in the research and clinical literature as “late talkers” (Paul, 1991; Paul & Alforde, 1993; Paul, Spangle-Looney, & Dahn, 1991; Rescorla, 1993; Rescorla & Schwartz, 1990; Thal et al., 1991; Thal & Tobias, 1992, 1994). Furthermore, studies suggest that approximately half of the children who experience such delays at 18–24 months of age continue to have language-learning problems in the late preschool years (Fischel, Whitehurst, Caulfield, & DeBaryshe, 1988; Paul, 1991; Paul & Alforde, 1993; Rescorla, 1993; Thal et al., 1991; Thal & Tobias, 1992). A central issue then becomes specifying factors that predict persistent versus transient language difficulties (i.e., that provide prognostic information about language outcomes).

Predicting outcome also is pertinent in language intervention. More specifically, it is beneficial to know which profiles of language impairment are likely to be the most responsive to intervention and in what ways. This information is important to families that are working toward understanding language impairment and striving to make future plans for their children, as well as for clinicians who are
seeking to develop optimal intervention plans for the present and future needs of the children that they serve.

Research findings suggest that measures of vocabulary comprehension may be useful in predicting development of students' language abilities and estimating intervention outcomes. First, in differentiating the overall population of late talkers into those with persistent versus transient delays, Thal et al. (1991) found that two measures of language comprehension are strong predictors of developmental outcome. In this project, 10 late talkers were identified at roughly 2 years of age and evaluated again 1 year later. At the 1-year follow-up evaluation, four of the children still experienced delays (true delays), whereas six demonstrated age-appropriate language skills (late bloomers). Examining data from the first visit, Thal et al. found that all of the children with true delays from the first visit were equivalent to those of an age-matched control group.

Second, Rice and Hadley (1995) reported language outcomes of children enrolled in a preschool classroom–based intervention program (Language Acquisition Preschool [LAP], Rice & Wilcox, 1995). Thirty-six children with SLI were enrolled in the intervention program for either 1 or 2 academic years. The protocol used to monitor progress and evaluate intervention efficacy included the PPVT–R, the Reynell Developmental Language Scales (Reynell & Gruber, 1990), and mean length of utterance (MLU). Overall, children with SLI gained an average of 9–11 standard score points on the language measures following enrollment in the intervention program. These standard score gains indicate that the children with SLI were learning new language skills at a rate that exceeded typical developmental expectations.

Of interest here are the observed profiles of language impairment among the children studied and their associations with outcome measures. Rice and Hadley (1995) found that children with an expressive disability profile at the outset of intervention programming (i.e., delays in expressive language and speech but age-appropriate receptive language, of which PPVT–R was a key component) made dramatic gains in expressive language skills during intervention (i.e., increases of 14–18 standard score points) and demonstrated average-range abilities on all language measures at preschool exit. Children with global language disabilities at the time of enrollment in the program (receptive, expressive, and speech delays) displayed standard score gains of 5–12 points but remained below the average range on all language measures except the PPVT–R at preschool exit. In addition, 79% of the children with global language impairments continued to receive intervention services through kindergarten and early elementary grades. In contrast, 56% of the expressive group were enrolled in language intervention in kindergarten; enrollment declined to 33% by the end of second grade. In summary, Rice and Hadley (1995) reported that for children with SLI enrolled in the LAP intervention program, intact receptive language abilities were associated with substantive gains in expressive language abilities and a reduced rate of future enrollment in intervention programming, relative to children with global disabilities.

Research findings in these two areas converge to suggest a predictive capacity for measures of vocabulary comprehension. Relatively little is known about the overall strength or confines of relations between early vocabulary comprehension and long-term language development status and/or responsiveness to early intervention. Important issues are whether vocabulary knowledge is the most robust predictor of future linguistic competence and the extent to which productive language difficulties are linked to later outcomes. In an investigation similar to that of Thal and her colleagues (Thal et al., 1991), Paul et al. (1991) did not find a pattern of greater risk for continued delay among children with concomitant receptive and expressive impairments. The discrepancy between studies may be attributed to differences in the particular measures of language competence that were used; Thal et al. (1991) specifically measured vocabulary comprehension, whereas Paul et al. (1991) measured general receptive and expressive language skills through the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984). Additional investigation is needed to sort out the particulars of links between early vocabulary knowledge and later language outcomes. (See Chapter 1 for further discussion.)

**Related Domains**

Research has revealed that children with language disabilities are at risk for difficulties in social interaction and acceptance and may be vulnerable to negative social bias (cf. Craig & Washington, 1993; Fujiki & Brinton, 1994; Hadley & Rice, 1991; Rice, 1993; Rice, Hadley, & Alexander, 1993; Rice, Sell, & Hadley, 1991; Windsor, 1995). An investigation by Gertner et al. (1994) assessed positive and negative peer nominations among three groups of preschoolers who attended a classroom-based language intervention program: children with typically developing language skills, children with speech and/or language impairments (S/LI), and children learning English as a second language. The peer nomination task asked children to “point to the picture of who you like to play with” and “point to the picture of who you do not like to play with,” with three positive and three negative nominations obtained from each child.

Multiple regression analyses revealed that 1) PPVT–R standard scores were the single best predictor of positive peer nominations, and 2) age and IQ did not significantly add to the variance in positive peer nominations accounted for by language measures (PPVT–R, Reynell Developmental Language Scales). Furthermore, examination of the patterns of peers’ nominations for individual children with S/LI revealed that language comprehension skills (as measured by performance on the PPVT–R and Reynell Receptive Scale) were a key factor.
in social acceptance. Children with S/LI whose receptive language skills were age appropriate received fewer-than-average negative nominations. In contrast, children whose S/LI profile included receptive language impairments tended to receive more negative peer nominations.

**Intervention Studies**

Measures of vocabulary comprehension also have played a central role in numerous treatment efficacy studies. The Rice and Hadley (1995) investigation of children’s PPVT-R standard score gains following enrollment in the Language Acquisition Preschool is an example of this use of the measure. Standardized measures can be optimal tools for assessing generalized treatment outcomes. Bearing in mind restrictions on the frequency of administering standardized tests, using vocabulary comprehension indices as treatment outcome measures enables the evaluation of global language change, supplementing the more common and circumscribed focus on specified treatment targets. Furthermore, using standard scores permits evaluation of whether language change exceeds that anticipated on the basis of maturation alone.

Research evaluating the influence of enhancing parental book-reading skills on children’s rate of language learning is one line of inquiry that has heavily used standardized vocabulary tests such as the PPVT-R as intervention outcome measures. In a series of studies, Whitehurst and his colleagues (Arnold et al., 1994; Whitehurst et al., 1988; Whitehurst et al., 1994) trained parents to use interactive strategies likely to enhance parent-child book reading; that is, strategies that provide optimal language models for children and encourage child participation in book-reading events (e.g., asking open-ended rather than yes-no questions, positively acknowledging child comments and questions, expanding on child comments). In these studies, the children had not been identified as having language impairment. In general, the design involves both experimental (parents learn about book-reading strategies) and control groups (no parent training occurs). The basic design also includes 1) a set of pretest measures of child language competence, 2) training parents in book-reading strategies and allowing a brief intervention period (roughly 6 weeks), and 3) readministering the measures of child language competence as a posttest. Pretest and posttest measures included the PPVT-R, as well as other standardized tests, such as the Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT-R) (Gardner, 1990) and particular subtests of the Illinois Test of Psycholinguistic Abilities (ITPA) (Kirk, McCarthy, & Kirk, 1968). In some of the studies, nonstandardized measures such as MLU and a specific test of expressive vocabulary designed by the investigators also were used as pretest and posttest measures of child language skill (cf. Whitehurst et al., 1994).

Findings of these investigations generally revealed positive influences of enhancing parental book-reading styles on the rate of children’s language acquisition, specifically in expressive language areas, and, to a lesser extent, in vocabulary comprehension (as measured by the PPVT-R). For example, Whitehurst et al. (1988) found that children in the experimental group scored significantly higher at posttest on the EOWPVT-R, the ITPA verbal expressive subtest, and MLU than did the control-group children (groups were equated at the outset of intervention). Posttest scores on the PPVT-R did not differ between groups. Valdez-Menchaca and Whitehurst (1992) implemented the basic design, outlined previously, in the context of Mexican child care; the experimental group scored significantly higher than did the control group at posttest on Spanish versions of the PPVT-R, the EOWPVT-R, and the ITPA. Arnold et al. (1994) contrasted a control group and two experimental groups, one in which parents received book-reading training through a videotape format and one in which parents received training in a typical passive instruction format. Posttest results revealed significant differences in PPVT-R standard scores between children whose parents received book-reading training in the videotape format versus children in the control and typical instruction groups.

In summary, measures of vocabulary comprehension have played a significant role in evaluating the effectiveness of various intervention programs and, more specific, have made a contribution to understanding how environmental manipulations in parental book-reading strategy and style can enhance the rate of children’s language development. Although expressive vocabulary skills appear more amenable to change, the Arnold et al. (1994) and Valdez-Menchaca and Whitehurst (1992) studies suggest that vocabulary comprehension abilities also can be enhanced through book-reading interventions. It is important to remember that the participants in these investigations are children with a full set of language-learning resources who are growing up in low-income environments. It may be the case that comprehension abilities are less easily promoted by brief enriching experiences, whereas production abilities, when language comprehension skills are intact, are more readily enhanced. An additional point of interest is the extent to which these changes are durable. Arnold et al. (1994) reported that investigations of the long-term consequences of such interventions for oral language development and emerging literacy development are under way. Researchers anticipate that measures of vocabulary comprehension will constitute a central component of the research protocol in assessing these issues.

**INTERPRETIVE BOUNDARIES:**

**LIMITATIONS OF VOCABULARY COMPREHENSION MEASURES**

The findings reviewed in the previous section point to the ubiquity of measures of vocabulary comprehension, particularly the PPVT-R. Furthermore, measures of vocabulary comprehension clearly are informative indices of linguistic ability and language change across a range of purposes and contexts. Recent studies, however, also have uncovered some surprising dissociations between vocabulary comprehension and other linguistic abilities. In addition, limitations in the popu-
Dissociations Between Initial and Full Mapping of Words

Several investigations have demonstrated that existing lexical knowledge (as measured by the PPVT-R) is not a powerful predictor of fast-mapping aptitude or QUIL abilities (cf. Rice, Buhr, et al., 1990; Rice et al., 1992). On the contrary, Rice and her colleagues have found little association between children's past lexical accomplishments (i.e., their accumulated lexicons) and their aptitude for gaining new words in comprehension during fast-mapping tasks. Initial-word mappings are only a tentative first step in the process of full mastery of lexical items. Although fast mapping is a very useful ability and one that almost certainly plays a central role in learning new words, the partial, or perhaps even superficial, word knowledge that begins in fast mapping must undergo substantial refinement before the novel lexical item is "acquired" in any complete sense. For example, a new form introduced through fast mapping must be associated with the existing lexicon and appropriately placed within the network of semantically related forms, the boundaries of the new word's meaning must be identified (i.e., the range of exemplars to which the word applies will need to be specified), and the grammatical function and role of the new form must be recognized. At best, competence in fully working through these steps seems to be indirectly, rather than directly, linked to initial-word-mapping aptitude (Rice, Buhr, et al., 1990; Rice et al., 1992).

Predicting Early Reading Success

Although vocabulary comprehension abilities are associated with literacy accomplishments (i.e., PPVT-R scores are significantly correlated with phonological, print, and metalinguistic awareness abilities [cf. Chaney, 1994; Tunmer et al., 1988]), vocabulary knowledge is only one contributor in a complex interaction of variables that influence reading achievement. For example, Mason (1992) identified multiple factors that influence reading acquisition, including home literacy experiences, language understanding and expression, and early decoding and print-labeling skills. Mason developed an intricate model to depict the interdependence of such factors in directing reading outcomes. These findings are substantiated by numerous investigations (see Dunning et al., 1994; Lonigan, 1994; Scarborough & Dobrich, 1994, for reviews).

Furthermore, vocabulary comprehension ability does not appear to be the most powerful predictor of the basic decoding abilities that are crucial for early reading achievement. In a longitudinal investigation of associations between spoken language development and reading acquisition, Catts (1993) and his colleagues (Catts, Hu, Larrivee, & Swank, 1994) assessed the oral language skills of preschoolers with and without language impairments, using both standardized measures of language skill (including the PPVT-R) and informal measures of phonological awareness (e.g., segmentation and blending tasks, rapid-naming measures). Results of multiple regression analyses revealed that the best predictors of early reading achievement were the phonological awareness measures. In a study of precocious language learners, Crain-Thoreson and Dale (1992) found that children with advanced linguistic and vocabulary abilities were not particularly facile early readers. In general, these findings concur with a number of other investigations demonstrating that performance on phonological awareness tasks is highly predictive of early reading achievement—more predictive than vocabulary abilities (Adams, 1990; Menyuk et al., 1990; Wagner & Torgesen, 1987).

Limitations with Diverse Populations

A significant issue in language assessment is identifying and developing measures that are appropriate for the diverse population of children whose language skills are evaluated. The widespread use of vocabulary comprehension measures such as the PPVT-R does not ensure that the measures are appropriate for the racially, ethnically, and economically diverse children who need appropriate assessment and intervention services or who may participate in research activities.

In fact, research findings suggest that, in particular circumstances, frequently used standardized measures are not appropriate. For example, Washington and Craig (1992) examined the performance on the PPVT-R of 105 urban, African American children from low-income backgrounds. The children were 4- to 6-year-olds enrolled in preschool or kindergarten programs at the time of the study. The mean standard score equivalent achieved on the PPVT-R was 79.7; nearly all children in this sample scored below the mean, and 65% scored below -1.00 standard deviation. A follow-up analysis revealed largely random error patterns, suggesting general bias in the overall test rather than specific content bias in a limited number of individual test items. Washington and Craig concluded that the PPVT-R is an inappropriate measure of vocabulary comprehension for African American children from urban, low-income environments. In brief, the measure provided limited information about individual abilities in vocabulary comprehension. The overall vocabulary tested by the PPVT-R and that known by the children appeared to be mismatched. In addition to potential bias present in the PPVT-R, the test context itself may have been alien to the children. Overall, the PPVT-R failed to provide useful information about the vocabulary knowledge of these participants. Other investigators also have suggested that the PPVT-R is an inappropriate measure for speakers of African American dialect (cf. Kreschek & Nicolosi, 1973).

Limitations with Very Young Children

As mentioned previously, one of the attractive features of vocabulary comprehension measures is their applicability across a broad age range; however, this advantage does not robustly apply to very young children. Standardized mea-
sures, such as the PPVT-R, generally require a reasonable amount of attention, ability to scan several pictures, and impulse control in selecting the frame named by the examiner. Very young children frequently find these task demands excessive. Norms for the PPVT-R begin at 2;6. Many children between the ages of 2;6 and 3.0 are capable of the task demands required by the PPVT-R, yet certainly some are not; few children below age 2;6 can complete a PPVT-R or other structured vocabulary comprehension tasks. Clearly, for children below age 3, standardized measures of vocabulary comprehension should be used with caution and interpreted in light of the child's interest in and attention to the task. With early identification and intervention services extending to the birth-to-3 period, the need for accurate and appropriate measures of language proficiency in this population is particularly pressing. However, measures of vocabulary comprehension in the picture-identification "show me X" format appear to be best-suited to children ages 2;6–3;0 and above. (See Chapter 8 for a discussion of using parent report in assessing vocabulary in young children.)

In summary, much of this chapter has been devoted to the insights gained from measures of vocabulary comprehension. This discussion would be both incomplete and inaccurate without acknowledging interpretive boundaries for measures of vocabulary comprehension, which include areas of dissociations between vocabulary knowledge and other developmental areas and limitations in the application of particular vocabulary comprehension measures.

CONCLUSION

Overall, one of the most striking aspects of the familiar "show me X" vocabulary comprehension task is that it is relatively unique among the multitude of available assessment tools and approaches in both the scope of uses to which it has been applied and the shared research and clinical insight it offers. Assessing vocabulary knowledge appears to tap some fundamental aspect of linguistic skill near the core of language capabilities. Perhaps the most convincing support for this conclusion comes from the literature on children with limitations in vocabulary knowledge. Insofar as restricted understanding of vocabulary is closely associated with negative social status among peers, it seems to strike at the very center of what it means to be an individual with less-than-typical linguistic abilities. Similarly, as variations in vocabulary knowledge relate to different developmental trajectories and varied outcomes of language intervention, the integral contribution of lexical skills to overall linguistic aptitude and knowledge is revealed.

The literature reviewed in this chapter indicates that a fair amount of what is known about the mechanisms and processes of both typical and atypical lexical development has come about through measures of vocabulary comprehension. Available literature suggests a fundamental clinical role for measures of vocabulary knowledge, with potential applications to clinical decision making and intervention planning, yet vocabulary knowledge also dissociates in some surprising ways from other linguistic competencies: existing vocabulary does not accurately predict how readily new words can be mapped and does not provide direct insight into probable reading achievement. Furthermore, measures of vocabulary comprehension, particularly standardized instruments such as the PPVT-R, appear vulnerable to cultural bias and must be used cautiously with individuals who differ from normative populations. Another population for which caution is warranted is that of very young children, in the age range of 2;6–3;0 and below, who may not successfully manage the task demands.

Despite the insights gained from measures of vocabulary comprehension, a number of issues have yet to be resolved. Of particular importance are issues that arise regarding clinical concerns. For example, additional information on the role of vocabulary and/or general comprehension abilities in differentiating children with persistent versus transient language difficulties is needed, as is more study delineating the contribution of intact versus impaired comprehension abilities to outcomes of language intervention. In both of these areas, longitudinal analyses of children with varied language comprehension and production profiles, including assessment of performance in specifically linguistic and related domains (i.e., reading and academic achievement), would be highly informative. Available research offers relatively broad clinical "hints" for anticipating outcomes (cf. Rice & Hadley, 1995; Thal et al., 1991); additional studies are needed to specify the particulars of clinical decision making and prognosis that seek to make use of vocabulary comprehension profiles.

In the general context of available means and methods of language assessment, this chapter conveys a relatively positive theme: Straightforward, brief, and widely used measures of vocabulary comprehension can be highly informative in research and clinical endeavors. Vocabulary comprehension tasks of the "show me X" type have contributed substantially to the knowledge base in child language development, disorders, and assessment.

REFERENCES


