Published in final edited form as:

Sch Psychol Q. 2009 September 1; 24(3): 147–150. doi:10.1037/a0017191.

Myth and Reality of the Word Caller: The Relation Between Teacher Nominations and Prevalence Among Elementary School Children

Elizabeth B. Meisinger,

Department of Psychology, University of Memphis

Barbara A. Bradley,

Department of Teaching and Leadership, University of Kansas

Paula J. Schwanenflugel,

Department of Educational Psychology and Instructional Technology, University of Georgia

Melanie R. Kuhn, and

School of Education, Boston University

Robin D. Morris

Department of Psychology, Georgia State University

Abstract

The purpose of the study was to investigate (a) the prevalence of word callers in elementary school, (b) the accuracy of teachers' word caller nominations, and (c) teachers' conceptualization of reading fluency and reading comprehension. To this end, 2 cross-sectional studies of second- and third- ($N = 868$) and of third- and fifth-grade ($N = 202$) children were conducted. Our findings suggest that word callers occur infrequently in the primary grades but that they are more prevalent in late elementary school. Regardless of grade level, teachers often overnominated children as word callers. Furthermore, a great deal of ambiguity and inconsistency seems to exist regarding teachers' understanding and use of the term. These findings suggest that the term should be used relatively rarely and that reading educators should be cautious about their identification of word callers in early elementary school.

Keywords

word caller; reading fluency; reading comprehension; teacher judgments

According to the National Reading Panel (2000), reading fluency is the ability to read text quickly, accurately, and with appropriate expression. It signifies the transition that students make as they become independent readers. The development of proficient reading fluency skills is a primary educational goal for elementary school children, and fluency plays an important role in reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Pikulski & Chard, 2005), the primary purpose of reading. There are numerous resources for teachers and reading specialists related to strategies that support reading fluency (e.g., Kuhn &

Correspondence concerning this article should be addressed to Elizabeth B. Meisinger, Department of Psychology, University of Memphis, 400 Innovation Drive, 202 Psychology Building, Memphis, TN 38152. bmsinger@memphis.edu.

Portions of this research were presented at the National Reading Conference in December 2004 in San Antonio, TX, and in December 2006 in Los Angeles, CA.

Schwanenflugel, 2007; Rasinski, Blachowicz, & Lems, 2006). However, some educators express concern that focusing on fluency-oriented practices will result in the creation of word callers, despite advice that fluency instruction should augment rather than replace comprehension instruction (Schwanenflugel & Ruston, 2007). Similar concern exists that the use of oral reading fluency probes in curriculum-based measurement (CBM) may overlook word callers (Hamilton & Shinn, 2003; Shapiro, 2004).

Word callers are children who efficiently decode words but do so without comparable comprehension taking place (Stanovich, 1986), so that words are called out without an understanding of the meaning of the text. Word callers are assumed to develop because there is an over-emphasis on pedagogical strategies that focus on basic reading skills instruction such as phonics (Nathan & Stanovich, 1991; Stanovich, 1986, 2000) and fluency (Hamilton & Shinn, 2003) rather than comprehension. However, a balanced instructional approach (i.e., one that focuses on decoding and fluency at the appropriate stages along with an emphasis on comprehension) is often utilized in today's classroom and is supported by the reading education literature (Pressley, Roehrig, Bogner, Raphael, & Dolezal, 2002). Furthermore, both word reading (Gough, Hoover, & Peterson, 1996; Shankweiler et al., 1999; Stanovich, 1980) and oral reading fluency (Fuchs et al., 2001; Jenkins, Fuchs, van den Broek, Epsin, & Deno, 2003b; Shinn, Good, Knutson, & Tilly, 1992) are closely related to reading comprehension.

The validity of the word caller phenomenon has been questioned by researchers (Hamilton & Shinn, 2003; Nathan & Stanovich, 1991; Stanovich, 1986, 2000), and has been described as a "red herring" of the reading literature (Nathan & Stanovich, 1991, p. 177). Despite the frequent use of the term (e.g., Walczyk & Griffin-Ross, 2007), there is almost no research to indicate what the actual prevalence of word calling among developing readers is and what teachers' understanding of the term is.

Conditions Related to Word Calling

Several conditions exist under which children may display a reading profile that is similar to that of a word caller. For example, hyperlexic readers are children who exhibit precocious development of word reading and decoding skills that is not accompanied by the development of comparable reading comprehension skills (Grigorenko, Klin, Pauls, Senft, Hooper, & Volkmar, 2002; Nation, Clarke, Wright, & Williams, 2006; Newman, Macomber, Naples, Babitz, Volkmar, & Grigorenko, 2007). Hyperlexia is associated with autism spectrum disorders (ASD) and is estimated to occur in 5–10% of the ASD population (Burd, Kerbeshian, & Fisher, 1985). Hyperlexia is thought to be part of a neurological disorder (ASD) that impairs children's social interactions, including their ability to appropriately use and understand language (Grigorenko et al., 2002; Nation et al., 2006; Newman et al., 2007). Therefore, hyperlexia is not the result of an overreliance on a particular pedagogical strategy; rather, it is part of a complicated array of symptoms caused by a neurological disorder manifesting in an inability to understand the social world.

Other children exhibit specific deficits in reading comprehension, meaning that they possess adequate word-level reading skills but struggle to understand what they read (Catts & Hogan, 2003; Shankweiler et al., 1999). Such specific deficits in reading comprehension are thought to be caused by broader impairments in language comprehension (Catts & Hogan, 2003; Stothard & Hulme, 1992). In other words, children with poor comprehension but adequate decoding skills have both lower verbal IQs and listening comprehension when compared with their normally comprehending peers but possess the normal phonological processing thought to support word reading skills (Stothard & Hulme, 1992). Furthermore, English-language learners may struggle with comprehension, even after they have successfully decoded the text, because of weaknesses in their English language skills (Carlo et al., 2004; Nakamoto, Lindsey,

& Manis, 2007; Proctor, Carlo, August, & Snow, 2005). In summary, the literature has identified some children with specific comprehension difficulties compared with their basic reading skills.

Who are Word Callers?

To our knowledge, only one study to date has examined the word caller phenomenon in any detail, and findings from that work lend little support to the validity of this construct. Hamilton and Shinn (2003) asked third-grade teachers whether they taught a student who could read fluently but had difficulty comprehending text. Of the 75 teachers surveyed, 31 indicated that they taught a student who matched that description, and 33 students were identified as word callers. Teachers were also asked to identify a peer who read with similar fluency as the word caller but with comprehension. Oral reading fluency skills were assessed with a curriculum-based measure (R-CBM) reading probe using a grade level passage. Reading comprehension was assessed using a CBM maze task, a comprehension oral question-answering test, and the Reading Comprehension subtest of the Woodcock–Johnson Reading Mastery Test (Woodcock, 1998; see also Hamilton & Shinn, 2003). It was expected that teacher-identified word callers would show fluency comparable with their peers but that word callers would comprehend less of what was read. However, teacher-identified word callers had weaker skills in both reading fluently *and* reading comprehension compared with their peers. That is, although teachers perceived all the students they selected for the study to have similar oral reading fluency, teacher-identified word callers read significantly and considerably more slowly than their peers. Moreover, despite using various criteria, Hamilton and Shinn were not able to identify any word callers in their sample. It is interesting that despite difference in rate, both groups read the connected text at a high level of accuracy (above 95%).

Hamilton and Shinn (2003) also compared teachers' judgments of students' oral reading fluency and reading comprehension skills to the students' actual performance by presenting teachers with the test materials and asking them to predict performance on each task (i.e., How many words will this student read in 1 minute? How many comprehension questions will this student answer correctly?). Teachers were found to overestimate both groups of students' reading skills on all measures. It is unclear to us whether teachers really know what typical reading rates and comprehension levels are, so this method may not be a valid indicator of teacher knowledge of students' reading skills.

Nevertheless, Hamilton and Shinn's findings call into question the accuracy with which teachers' identify children as word callers. It may be that teachers' conceptualization of oral reading fluency and comprehension differ from researcher-chosen outcome measures. Definitions of oral reading fluency do not usually include reading comprehension as an essential part, although some do conceptualize fluency as a kind of rate-limited comprehension process (Fuchs et al., 2001; Samuels, 2006; Wolf & Katzir-Cohen, 2001). The Hamilton and Shinn study was limited by the fact that only a small sample of students ($N = 66$) were tested, so the actual base rate of word calling is unclear. Furthermore, because only third graders were examined, it is unclear whether actual base rates of word calling might change as children become generally more fluent.

Purpose

The purpose of this research is to fill important gaps in the literature regarding base rates of word calling and the accuracy of teachers' judgments regarding word callers. Specifically, we hope to address the issue of the prevalence of word callers better by using a large sample to evaluate the prevalence of word callers in elementary school. Furthermore, we wanted to determine whether the incidence of word callers is stable throughout elementary school by examining its prevalence cross-sectionally. Schwanenflugel, Meisinger, Weisenbaker, Kuhn,

Strauss, and Morris (2006) found that the correlation between reading fluency and reading comprehension skills declines considerably as children get older, providing greater potential for word calling to emerge developmentally. Finally, we wanted to explore teachers' concepts of reading fluency and comprehension to determine whether these concepts can be related to their nominations of word callers.

Study 1

The aim of Study 1 was to investigate the prevalence of word callers in early elementary school-aged children. Children were assessed with standardized assessments of reading fluency and reading comprehension. Two criteria (one comparable with those used by other researchers and another inspired by the hyperlexia literature) were then applied to children's reading scores to determine the prevalence of identified word callers.

Method

Participants—Participants were 521 second-grade (M age = 8 years, 2 months; range = 7 years, 1 month to 9 years, 10 months) and 347 third-grade (M age = 9 years 4 months; range = 8 years 5 months to 10 years 9 months [N = 868]) students from 11 southeastern (n = 582) and northeastern elementary schools (n = 286) tested during the spring term. Of the children, 49.5% were African American, 17.1% were Caucasian, 25.9% were Hispanic, 5.5% were Asian, and 2.0% identified as other or were unknown; 48% were female, and 52% were male. The children attended schools in which 40%–95% of students received free or reduced-cost lunch. All attended general education classes, and none were excluded on the basis of reading disability or other special education eligibility unless they were receiving services in a self-contained as opposed to a mainstream classroom. Participants in Study 1 and 2 were part of a larger project examining the development of reading fluency. However, no overlap exists in participants of the two studies.

Reading Assessments—Measures were standardized assessments of both reading fluency and reading comprehension, as described here.

(a) Oral reading fluency: The Gray Oral Reading Test—Fourth Edition (GORT–4; Wiederholt & Bryant, 2001) is designed to assess oral reading fluency. It consists of a series of increasingly difficult passages, in which an examiner records the time in seconds for passage completion along with reading errors made. Children continue reading passages of increasing difficulty until they become too slow and inaccurate on the basis of the discontinue rule (i.e., a combination of reading speed and the number of misread words) set by the test developer. This test produces a rate, accuracy, and fluency score that combines the two. The GORT–4 scaled scores have a mean of 10 (SD = 3), which we converted into standard scores with a mean of 100 (SD = 15) to facilitate comparisons with other tests. The GORT–4 manual reported strong reliability evidence ranging from .90 to .96. Concurrent validity estimates with similar measures ranged from .39 to .89.

(b) Reading comprehension: The Wechsler Individual Achievement Test—Reading Comprehension subtest (WIAT–RC; Wechsler, 1992) was used to assess children's comprehension skills. The WIAT–RC subtest was chosen because it includes a key indicator of comprehension (i.e., the child can answer questions about text; Byers, 1998; Richardson, Anders, Tidwell, & Lloyd, 1991) and avoids false test floors that include items outside of the domain of comprehension (decoding, letter recognition, etc.). The subtest consists of several passages of increasing complexity that children read and answer questions about until they miss four sequential questions. The WIAT–RC generates a standard score with a mean of 100 (SD = 15). The WIAT manual reports reliabilities ranging from .85 to .99, and validity estimates

with other similar reading measures ranged between .74 and .79 (Wechsler, 1992). Research suggests that it correlates well with other tests of comprehension and that teachers view it as valid (Byers, 1998; Foegen, Espin, Allinder, & Markell, 2001; Smith & Smith, 1998).

Procedure—After obtaining parental consent, we explained the study to the children, and each child assented before testing. All assessments were administered individually in a quiet area of the school. Assessments were counterbalanced and required approximately 45 min to administer. Children received a sticker or pencil, and their teachers received books for their classroom as a token of thanks. A doctoral-level school psychology student who had substantial experience administering reading fluency and comprehension assessments trained all examiners. All examiners were trained to have at least 95% agreement with the trainer, and the trainer observed the examiners for the first week of testing to ensure procedural adherence.

Word Caller Identification Criteria—To ensure generality of findings over variations in the definition of word caller, we used two criteria to identify children as word callers. Criterion 1 is comparable with the definition used by other researchers (i.e., Hamilton & Shinn, 2003), which required that children possess psychometrically normal fluency skills (as indicated by a minimum standard score of 95 on the GORT-4 Fluency scale) in conjunction with below average comprehension skills (as indicated by a standard score of below 85 on the WIAT-RC). Therefore, a minimal difference of 10 standard score points (or $2/3$ of 1 *SD* on the test) was required to be identified as a word caller. The use of a buffer zone is thought to “mitigate some of the arbitrariness of a cut score” approach (Shankweiler et al., 1999, p. 75) and has been used in other research when dividing children into groups on the basis of their reading profile (e.g., Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003a).

Criterion 2, inspired by the literature on children with hyperlexia, required children to have strong oral reading skills (standard scores of 110 or higher, or $2/3$ *SD* above the test mean) but allowed for somewhat less impaired reading comprehension (standard scores of 90 or below, or $2/3$ *SD* below the test mean). The purpose of this second criterion was to provide a definition that captured children who had obviously excellent fluency but for whom there was a large discrepancy ($1\ 1/3$ *SD*) between their fluency and comprehension skills.

Results

Means for the reading achievement measures fell in the average range (WIAT-RC, $M = 99.44$, $SD = 12.88$; GORT-4 Fluency, $M = 95.0$, $SD = 17.28$; GORT-4 Accuracy, $M = 95.35$, $SD = 17.36$; GORT-4 Rate, $M = 96.85$, $SD = 16.60$). The children’s data were classified according to the two criteria used to identify word callers (see Table 1). When Criterion 1 was applied to the GORT-4 Fluency and WIAT-RC subtest scores, 1.4% ($n = 12$) of the total sample could be identified as word callers (0.8% were second-grade children, and 2.3% were third-grade children). When Criterion 2 was applied, only 0.8% ($n = 7$) of the total sample were identified as word callers (0.4% were second-grade children, and 1.4% were third-grade children). There was no relationship between the prevalence of word callers and grade, $\chi^2(1, N = 868) < 3.84$, $p > .05$.

Hamilton and Shinn’s (2003) findings suggested that teachers may identify slow but accurate, rather than fluent, readers as word callers. To explore this possibility, we applied Criteria 1 and 2 to the GORT-4 Accuracy subtest and the WIAT-RC subtest. In this manner, word callers could be identified if they read accurately but with poor comprehension. Results from the application of Criterion 1 on reading accuracy were comparable with those for fluency, with 1.2% ($n = 10$) of the sample identified as a word caller (1.2% were second-grade children, 1.2% third-grade children). Similarly, using Criterion 2 on reading accuracy, 2.1% ($n = 18$) of children could be identified as word callers (1.0% second-grade children, 3.7% third-grade

children). Again, no relationship was found between prevalence and grade level, $\chi^2(1, N = 868) < 3.84, p > .05$. In summary, Study 1 suggests that word callers do not exist in appreciable numbers among early elementary school children. The use of accuracy or rate, and varying criteria, produced similar results.

Study 2

Hamilton and Shinn's (2003) research called into question the accuracy of teachers' word caller nominations and judgments of children's reading skills. However, it is unclear whether their findings could be attributed to the method used to assess teachers' judgments of reading skills or the fact that they examined only a single grade level. Thus, one purpose of Study 2 was to examine the accuracy of teachers' word caller nominations. To this end, students' reading fluency and reading comprehension skills were assessed by their teachers both through ratings and through standardized measures. A second purpose of Study 2 was to understand teachers' concepts regarding skills underlying word calling. Variation in their concepts was examined in relation to the skills of the children that they nominated as word callers. Last, Study 2 examined changes in word caller prevalence as children approached the end of elementary school.

Method

Participants—Participants were 110 third-grade students (M age = 9 years, 4 months; range = 8 years, 7 months to 10 years, 10 months) and 92 fifth-grade students (M age = 11 years, 4 months; range = 10 years, 5 months to 12 years, 6 months; $N = 202$) from five schools located in the southeastern United States tested in the spring term. Of the children, 51.5% were classified as African American, 38.6% as Caucasian, 6.9% as Hispanic, and 3.0% were classified as other or unknown; 56.4% were female and 43.6% were male. Approximately 77% received free or reduced-cost lunch. All attended general education classes and were not excluded on the basis of reading disability or other special education eligibility unless they received services in a self-contained classroom. Twenty-one teachers participated in the study (14.3% African American; 85.7% Caucasian; 81.0% female, 19.0% male).

Teacher Assessments—Teachers completed several questionnaires pertinent to the topic of word calling.

(a) Teacher questionnaire: Teachers were administered a questionnaire and asked to define the terms *reading fluency* and *reading comprehension*. Specifically, teachers were asked, "In your own words, how would you define reading comprehension?" and "In your own words, how would you define reading fluency?" Teachers' written responses to the reading fluency question were then classified as having either a *basic reading fluency* or an *expanded reading fluency* definition. Teacher responses were classified as a *basic reading fluency* definition when at least one of the three components of fluency (i.e., speed, accuracy, & appropriate expression; National Reading Panel, 2000) was described. Responses were labeled as an *expanded reading fluency* definition if they also included comprehension processes (e.g., make connections within the text, monitor comprehension, ability to summarize the story). Teachers' definitions of reading comprehension were classified as a *basic reading comprehension* definition if only comprehension processes were mentioned or an *expanded reading comprehension* definition if one or more aspects of fluency were mentioned in addition to comprehension. Two graduate students trained in this classification system scored teachers' responses. To train for reliability, Elizabeth B. Meisinger explained the coding system in detail, and the students practiced until they reached a 90% agreement with her. Interrater reliability was determined using a Cohen's kappa coefficient ($k = .90$) based on 20% of the questionnaires. Disagreements were infrequent and resolved through discussion.

(b) Word caller nominations: We presented teachers with a list of children for whom consent was granted, and we asked them to circle the names of those they would describe as word callers, defined as children “who can read fluently but have difficulty comprehending the text.”

(c) Teacher ratings of children’s reading skills: Teachers were provided these instructions for rating their students’ reading skills:

On the following page is a list of the children in your class who are participating in our study. Please rate each child’s reading fluency and comprehension using the two following scales. Assume that the children are reading from grade level text (please see attached example).

To rate comprehension, they were asked to use a 3-point scale on which 1 = *low comprehension skills* (i.e., the child struggles to understand what he/she reads such that he/she stands out among the other children as experiencing difficulty in comprehension), 2 = *average comprehension skills* (i.e., the child understands what he/she read about as well as most other children did), and 3 = *high comprehension skills* (i.e., the child comprehends better than most children his or her age). To rate fluency, we asked teachers to use the National Assessment of Education Progress (NAEP) Oral Fluency Scale (Pinnell, Pikulski, Wixson, Campbell, Gough, & Beatty, 1995), on which 1 = *primarily word-by-word reading with occasional 2- or 3-word phrases*; 2 = *reading characterized by primarily 2-word phrases with some 3- or 4-word phrases*; 3 = *reading primarily characterized by 3- and 4-word phrase groups with little expression*, and 4 = *fluent reading characterized by expressive reading in larger, meaningful phrases with few regressions, repetitions, or deviations*.

Procedure—We administered the same fluency and comprehension assessments to children as those in Study 1, using the same procedures. Teacher questionnaires were distributed in a single packet. The instructions for the three questionnaires were individually reviewed with each teacher and any questions they had were addressed. All teachers completed these documents independently and returned completed forms within 2 weeks of completing the child assessments.

Results

Standardized measures indicated that, on average, children possessed age-appropriate reading skills (WIAT-RC, $M = 98.91$, $SD = 12.97$; GORT-4 Fluency, $M = 95.60$, $SD = 17.94$). Teacher ratings of students’ skills also indicated that children demonstrated average reading comprehension (Comprehension scale, $M = 2.23$, $SD = 0.75$) and fluent oral reading skills (NAEP Oral Fluency Scale, $M = 3.11$, $SD = 0.90$).

The two criteria used in Study 1 were used to identify word callers. When Criterion 1 was applied to the data, a relationship was revealed between grade and word caller rate, $\chi^2(1, N = 202) = 6.17$, $p < .05$, so that a greater percentage of fifth-grade (9.78%) than third-grade children (1.82%) were identified as word callers. However, with Criterion 2, no relationship between grade and word caller rate was found, $\chi^2(1, N = 202) = 1.20$, $p > .05$; no third graders and only one fifth grader (1.1%) was identified. An examination of the correlations between reading fluency and comprehension revealed several interesting trends (see Table 2). Moderate to strong correlations were observed in each grade. However, a declining trend, consistent with Schwanenflugel et al.’s (2006) findings, was apparent between reading fluency and comprehension. In other words, these skills were highly related in younger students ($r = .72$, $p < .01$), but the relation decreased in older students ($r = .51$, $p < .01$), $z = 2.41$, $p < .05$. These findings suggested that word callers with psychometrically poor comprehension skills may be more prevalent among older students.

Accuracy of Teachers' Word Caller Nominations

The majority of teachers (71.4%) nominated at least one student as a word caller, and 22.3% of children were nominated by their teacher as a word caller (22.7% in third grade, and 21.7% in fifth grade). Of particular interest was the relation between researcher-identified and teacher-nominated word callers. Regardless of criteria applied, no association was found between researcher-identified and teacher word caller nominations, $\chi^2(1, N = 202) < 3.84, p > .05$. As shown in Table 3, using Criterion 1, teachers identified 75.2% of children correctly as either a word caller ($n = 3$) or nonword caller ($n = 149$), but 93.3% of teacher-nominated word callers were false-positives ($n = 42$). Furthermore, teachers failed to identify some students ($n = 8$) who did fit the criterion. Thus, teachers not only overidentified children as word callers but overlooked some actual word callers in their class.

Profile of Teacher-Nominated Word Callers

Children nominated by their teacher as a word caller would be expected to possess fluent reading skills but struggle to comprehend text. To investigate the accuracy of teachers' word caller nominations, we examined children's standardized test reading profiles to determine whether their reading skills would meet the expected profile. To this end, we conducted a series of one-way analyses of variance (ANOVAs) comparing comprehension, fluency, and accuracy scores as a function of teacher's word-caller nomination (i.e., word caller or nonword caller). Significant differences were observed across teacher-nominated word callers and their peers on comprehension, $F(1, 200) = 15.44, p < .001$, partial $\eta^2 = .07$; and fluency, $F(1, 200) = 4.49, p < .05$, partial $\eta^2 = .02$. Specifically, teacher-nominated word callers had lower comprehension scores and lower fluency scores, compared with their peers (see Table 4), but read connected text as accurately as their peers, $F(1, 200) = 1.83, p > .05$, partial $\eta^2 = .01$.

To assess whether teachers' concepts of fluency and comprehension differed from the standardized, norm-referenced tests used, we compared teachers' word caller nominations with their own ratings of students' fluency and comprehension. A series of one-way ANOVAs was conducted examining teacher word caller nominations (i.e., word caller or nonword caller) on teachers' comprehension and fluency ratings. As with the standardized, norm-referenced measures (see Table 4), teacher-nominated word callers were rated by their teacher as being less fluent, $F(1, 200) = 13.32, p < .001$, partial $\eta^2 = .06$; and having lower comprehension, $F(1, 200) = 43.53, p < .001$, partial $\eta^2 = .18$; compared with their peers (see Table 4). Thus, unlike the standard definition of word caller, these teachers appeared to be nominating children who had both poor fluency and poor comprehension.

Teachers' Definitions of Reading Skill

We examined teachers' definitions of fluency and comprehension to better understand how teachers conceptualize the skills underlying word calling. For the concept fluency, 38.1% of the teachers provided a *basic fluency* definition (i.e., they mentioned only features of rate, accuracy, and/or appropriate expression) and 61.9% provided an *expanded reading fluency* definition, including comprehension features in their definition. By contrast, when teachers' definitions of reading comprehension were examined, 90.5% described only comprehension features (e.g., make connections within the text, comprehension monitoring, ability to summarize or retell the story) with only 9.5% mentioning an aspect of reading fluency (i.e., rate, accuracy, appropriate expression). Thus, most teachers' perceive comprehension as a component of fluent reading, but not the reverse.

The concept of word caller is based on an incongruity between children's fluency and comprehension skills. Because many teachers conflate fluency and comprehension conceptually, it seemed likely that variation in teachers' conceptualization of fluency might affect the reading profile of students they nominated as word callers. Specifically, it was

hypothesized that teachers who conflate these terms would likely nominate children who were struggling with both comprehension and fluency as word callers, whereas teachers who conceptualize fluency and comprehension as distinct skills would be more likely to nominate children who met the standard profile of a word caller (i.e., fluent readers with poor comprehension). Therefore, a 2 (fluency definition: basic vs. expanded definition) $\times$ 2 (teacher-nominated word caller: word caller or peer) ANOVA was conducted for fluency and comprehension separately. No main effect was found for fluency definition on either comprehension, $F(1, 198) < 1.0, p > .05$, partial $\eta^2 < .00$; or fluency, $F(1, 198) < 1.0, p > .05$, partial $\eta^2 < .00$. Furthermore, the interaction between fluency definition and word caller nomination was not significant for either comprehension, $F(1, 198) = 1.63, p > .05$, partial $\eta^2 = .00$; or fluency, $F(1, 198) < 1.00, p > .05$, partial $\eta^2 = .00$. Teachers' concepts of fluency did not translate into observed differences in the reading profile of children they nominated as word callers.

Who Are Word Callers?

Our results suggest that teacher-nominated word callers do not meet the expected profile of a word caller based on either children's standardized assessments or teachers' own ratings of students' reading skill. Furthermore, the disconnection between children's reading profiles and teachers' word caller nominations was not explained by variation in teachers' definition of fluency. Because our previous analysis indicated that, on average, teacher-nominated word callers are children who read with less fluency and less comprehension than their peers, it seemed reasonable to posit that teachers' may simply be nominating poor or struggling readers. To test this hypothesis, we identified children with standard scores less than 90 on both fluency and comprehension measures and classified them as poor readers. These poor readers were then compared with teacher-nominated word callers in a contingency table (see Table 5) to examine whether there was overlap in these groups of children. However, only 24.4% of teacher-nominated word callers were also identified as poor readers. In summary, it does not appear that teachers are merely nominating poor readers as word callers, although some teachers do seem to endorse this conceptualization of the term, at least implicitly. It is important to note that our findings suggest that teachers believe that there are more word callers in their classrooms than there actually are.

General Discussion

Little research has examined the validity of the word caller concept, despite its potential relevance to pedagogy and literacy assessment. Concern for word callers is sometimes raised by reading educators regarding phonics or fluency-related instructional practices (Hamilton & Shinn, 2003; Nathan & Stanovich, 1991; Schwanenflugel & Ruston, 2007; Stanovich, 1986, 2000) and CBM reading assessments (i.e., oral reading fluency probes; Hamilton & Shinn, 2003; Shapiro, 2004). The present studies provided clarification regarding the prevalence of word callers among elementary school children, the accuracy of their teachers in discerning them, and teachers' concepts of fluency and comprehension that form the basis word caller nominations. Consistently, we found that word callers do not exist in appreciable numbers in early elementary school, regardless of the assessments or criteria used. However, by fifth grade, nearly 10% of students demonstrated normal reading fluency but below average comprehension skills.

The aforementioned finding supports the idea that there is a growing asymmetry between basic reading skills (e.g., word reading and fluency) and comprehension as children get older (Schwanenflugel et al., 2006; Vellutino, Fletcher, Snowling, & Scanlon, 2004). There are several likely reasons why fluency may become increasingly disconnected from comprehension. First, the more complex texts that children in the upper elementary grades are

expected to read demand skills other than fluency. In young readers, reading fluency and comprehension are thought to be closely related, partly because of the attentional resources that are made available as children develop automatic word recognition skills (LaBerge & Samuels, 1974; Perfetti, 1985). That is, once children are able to quickly and accurately identify the words they encounter in text, attentional resources can be reallocated to comprehension. At some point, however, these lower levels of automaticity are insufficient for good comprehension because the issue shifts from freed resources to issues of background knowledge, metacognition, and a myriad of other skills and strategies. Comprehension skills such as the utilization of content knowledge and sophisticated vocabulary, the ability to draw inferences and abstract relations from the text, and the ability to monitor comprehension become more important once children have mastered basic fluency skills (Chall, 1996; Sweet & Snow, 2003). Second, it may be that silent reading fluency is more related to comprehension in older students than oral reading is, as children have typically transitioned to this mode of reading by fourth or fifth grade (Prior & Welling, 2001). Perhaps fewer word callers would be identified if silent reading fluency measures had been used. Third, it may be that some elementary school teachers focus on comprehension skills more than others, the cumulative effect of which would unfold as children grow older. However, given the variety of schools, districts, and regions from which our participants came, we consider this to be less likely.

Findings from this study confirm previous research (i.e., Hamilton & Shinn, 2003) that questions the accuracy of teachers' word caller nominations. Nearly one quarter of young children in Study 2 were identified by their teacher as word callers, despite the relatively low incidence of true word callers (0–2.3%) in early elementary school. Although some true word callers were actually overlooked, the preponderance of teacher error was made in the direction of overnominating children as word callers. Instead, children identified by their teacher as word callers were, on average, less fluent and comprehended less of what was read on reading assessments. However, the number of children involved in the Hamilton and Shinn (2003) study was very small in comparison with this one, and no true assessment of the incidence of word calling could be made on the basis of that study because of the matched sample design they used. By contrast, the present study with its 1,070 children confirmed that the incidence of word calling, especially in early elementary school, is indeed small but increases as children get older.

It is unclear why some children were nominated by their teachers as word callers. The discrepancy between teacher-nominated and psychometrically identified word callers might be explainable by discrepancies between teachers' concepts of fluency and reading comprehension and those of the test manufacturers. However, teachers rated the children they nominated as having worse fluency *and* comprehension. Furthermore, teacher-nominated word callers read connected text as accurately as their peers, suggesting that, unlike the teachers in Hamilton and Shinn's (2003) study, teachers were not relying on reading accuracy for these nominations. Although it is possible that teachers were simply nominating poor readers (i.e., dysfluent readers with poor comprehension), standardized assessment analysis from Study 2 suggest that only about one quarter of the children nominated by their teachers were very poor readers, leaving approximately three quarters of teachers' nominations unexplained. In summary, results from this work substantiate the skepticism toward the accuracy of teachers' word caller nominations.

Next, we were able to identify some differences between teacher concepts of fluency skills and that of the test makers. More than half of the teachers in Study 2 considered comprehension to be a key component of fluent reading, whereas fluency was seldom seen as a necessary component of comprehension. It is possible that conflating fluency and comprehension conceptually might influence a teacher's word caller nominations. Unfortunately, this variation

in teachers' definition of fluency did not translate into differences in the reading profiles of children they nominated as word callers.

Finally, our findings have implications regarding the use of CBM reading assessments. For younger students, the strong relation between reading comprehension and fluency in emerging reading skills may make fluency a reasonable proxy for general reading skill (Fuchs et al., 2001). However, this research suggests that we should be concerned about overlooking word callers among older students. Reading fluency measures may need to be augmented with comprehension measures in older students. Indeed, other research suggests that the predictive accuracy of CBM reading probes may be improved by supplementing these assessments with a reading comprehension task, especially for late elementary school children (Schilling, Carlisle, Scott, & Zeng, 2007; Shapiro, Solari, & Petscher, 2008).

Limitations and Future Directions

Several limitations of this research warrant discussion. First, data from fourth-grade children were not available. Fourth grade is said to be marked by a transition from learning to read to reading to learn (Chall, 1996). Children's comprehension at this point often suffers from what has come to be known as the "fourth-grade slump" (Chall & Jacobs, 2003). Studying the development of this slump across the fourth-grade year may provide key insights into the development of word calling among some children. Next, this work should be replicated with a larger sample of later elementary and middle school students to strengthen the generalizability of the results and to continue to explore developmental variations in this phenomenon. Given that students transition to silent reading fluency by later elementary school, investigations involving older students should incorporate measures of silent reading fluency. Also, it could be argued that the use of norm-referenced tests limits this work, as our reading fluency assessment may not be interchangeable with R-CBM measures. However, our results regarding the prevalence of word callers and the accuracy of teachers' word caller nominations replicated that of Hamilton and Shinn (2003), who utilized both norm-referenced and curriculum-based measures, suggesting that our results were not due to the type of assessment used. Furthermore, teachers based their word caller nominations on the researcher's definition of the term. If teachers were allowed to use their own definition, perhaps less ambiguity would surround their word caller nominations. However, it is clear that teacher nominations did not conform to the definition provided them. Also, instructing teachers to nominate children as word callers may have set the expectation that they should have word callers in their class. Because 28.6% of teachers did not nominate any children as word callers, we believe that these expectations were not so strong that they completely overrode teachers' sound judgments. Still, future investigations should provide instructions that explicitly state that there may or may not be word callers in a class. Last, child behavioral data were not available; this precluded our examination of behavioral factors in teachers' word caller nominations. Future research should address these limitations, as well as explore the developmental trends between both oral and silent reading fluency and comprehension in older students.

Conclusions

In summary, our results suggest that word calling occurs infrequently among early developing readers but emerges as a discernable phenomenon toward the end of elementary school. However, teachers are not accurate at identifying children as word callers. Despite repeated attempts at discerning the conceptual basis for the term among teachers, the rationale for teachers' word caller nominations remains unclear. It is up to future researchers to help teachers better define the construct or to assist educators in the development of an alternative definition that better captures the reading behaviors they are trying to explain.

Acknowledgments

This project was supported by the Interagency Education Research Initiative, a program of research managed jointly by the National Science Foundation, the Institute of Education Sciences in the U.S. Department of Education, and the National Institute of Child Health and Human Development in the National Institutes of Health (NICHD NIH). Funding for the project was provided by NICHD NIH Grant 7 R01 HD040746-06.

References

- Burd L, Kerbeshian J, Fisher W. Inquiry into the incidence of hyperlexia in a state-wide population of children with pervasive developmental disorder. *Psychological Reports* 1985;57:236–238. [PubMed: 4048337]
- Byers JA. Content and concurrent validity of the WIAT and WJ-R reading subtests for second grade students. *Dissertation Abstracts International* 1998;5:3100.
- Carlo MS, August D, McLaughlin B, Snow CE, Dressler C, Lippman DN, ... & White CE. Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly* 2004;39:188–215.
- Catts HW, Hogan TP. Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology* 2003;24:223–246.
- Chall, JS. Stages of reading development. 2. Orlando, FL: Harcourt Brace; 1996.
- Chall JS, Jacobs VA. Poor children's fourth-grade slump. *American Educator* 2003;27:14–44.
- Foegen A, Espin CA, Allinder RM, Markell MA. Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education* 2001;34:226–236.
- Fuchs LS, Fuchs D, Hosp MK, Jenkins JR. Text fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading* 2001;5:239–256.
- Gough, PB.; Hoover, WA.; Peterson, CL. Some observations on a simple view of reading. Mahwah, NJ: Erlbaum; 1996.
- Grigorenko EL, Klin A, Pauls DL, Senft R, Hooper C, Volkmar F. A descriptive study of hyperlexia in a clinically referred sample of children with developmental delays. *Journal of Autism and Developmental Disorders* 2002;32:3–12. [PubMed: 11916330]
- Hamilton C, Shinn MR. Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review* 2003;32:228–240.
- Jenkins JR, Fuchs LS, van den Broek P, Espin C, Deno SL. Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice* 2003a;18:237–245.
- Jenkins JR, Fuchs LS, van den Broek P, Espin C, Deno SL. Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology* 2003b;95:719–729.
- Kuhn, MR.; Schwanenflugel, PJ. Fluency in the classroom. New York: Guilford Press; 2007.
- LaBerge D, Samuels S. Toward a theory of automatic information processing in reading. *Cognitive Psychology* 1974;6:293–323.
- Nakamoto J, Lindsey KA, Manis FR. A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing* 2007;20:691–719.
- Nathan RG, Stanovich KE. The causes and consequences of differences in reading fluency. *Theory Into Practice* 1991;30:176–184.
- Nation K, Clarke P, Wright B, Williams C. Patterns of reading ability in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 2006;36:911–919. [PubMed: 16897396]
- National Reading Panel. Report of the national reading panel. Washington DC: Author; 2000.
- Newman TM, Macomber D, Naples AJ, Babitz T, Volkmar F, Grigorenko EL. Hyperlexia in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2007;37:760–774. [PubMed: 17048093]
- Perfetti, C. Reading ability. New York: Oxford Press; 1985.
- Pikulski JJ, Chard DJ. Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher* 2005;58:510–519.

- Pinnell, GS.; Pikulski, JJ.; Wixson, KK.; Campbell, JR.; Gough, PB.; Beatty, AS. The Nation's Report Card. Report No. 23-FR-04. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.; 1995. Listening to children read aloud: Data from NAEP's integrated reading performance record (IRPR) at Grade 4.
- Pressley M, Roehrig A, Bogner K, Raphael LM, Dolezal S. Balanced literacy instruction. Focus on Exceptional Children 2002;34:1–14.
- Prior SM, Welling KA. "Read in your head": A Vygotskian analysis of the transition from oral to silent reading. Reading Psychology 2001;22:1–15.
- Proctor CP, Carlo M, August D, Snow C. Native Spanish-speaking children reading in English: Toward a model of comprehension. Journal of Educational Psychology 2005;97:246–256.
- Rasinski, T.; Blachowicz, C.; Lems, K. Fluency instruction: Research-based best practices. New York: Guilford Press; 2006.
- Richardson V, Anders P, Tidwell D, Lloyd C. The relationship between teachers' beliefs and practices in reading comprehension instruction. American Educational Research Journal 1991;28:559–586.
- Samuels J. Looking backward: Reflections on a career in reading. Journal of Literacy Research 2006;38:327–344.
- Schilling SG, Carlisle JF, Scott SE, Zeng J. Are fluency measures important predictors of reading achievement? The Elementary School Journal 2007;107:429–448.
- Schwanenflugel PJ, Meisinger EB, Wisenbaker JM, Kuhn MR, Strauss GP, Morris RD. Becoming a fluent and automatic reader in the early elementary school years. Reading Research Quarterly 2006;41:496–522. [PubMed: 20072665]
- Schwanenflugel, PJ.; Ruston, HP. Becoming a fluent reader: From theory to practice. In: Kuhn, MR.; Schwanenflugel, PJ., editors. Fluency in the classroom: Solving problems in the teaching of literacy. New York: Guilford Press; 2007. p. 1-16.
- Shankweiler D, Lundquist E, Katz L, Stuebing KK, Fletcher JM, Brady S, ... &, Shaywitz A. Comprehension and decoding: Patterns of association in children with reading difficulties. Scientific Studies of Reading 1999;3:69–94.
- Shapiro, ES. Academic skills problems: Direct assessment and intervention. 3. New York: Guilford Press; 2004.
- Shapiro ES, Solari E, Petcher Y. Use of a measure of reading comprehension to enhance prediction on the state high stakes assessment. Learning and Individual Differences 2008;18:316–328.
- Shinn MR, Good RH, Knutson N, Tilly WD. Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. School Psychology Review 1992;21:459–479.
- Smith TD, Smith BL. Relationship between the Wide Range Achievement Test 3 and the Wechsler Individual Achievement Test. Psychological Reports 1998;83:963–967. [PubMed: 9923176]
- Stanovich KE. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly 1980;16:32–71.
- Stanovich KE. Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly 1986;21:360–406.
- Stanovich, KE. Progress in understanding reading: Scientific foundations and new frontiers. New York: Guilford Press; 2000.
- Stothard SE, Hulme C. Reading comprehension difficulties in children: The role of language comprehension and working memory skills. Reading and Writing 1992;4:246–256.
- Sweet, AP.; Snow, CE. Rethinking reading comprehension: Solving problems in the teaching of literacy. New York: Guilford Press; 2003.
- Vellutino FR, Fletcher JM, Snowling MJ, Scanlon DM. Specific reading disability (dyslexia): What have we learned in the past four decades? Journal of Child Psychology and Psychiatry 2004;45:2–40. [PubMed: 14959801]
- Walczyk JJ, Griffin-Ross DA. How important is reading skill fluency for comprehension? The Reading Teacher 2007;60:560–569.
- Wechsler, D. Wechsler individual achievement test. San Antonio, TX: Psychological Corporation; 1992.
- Wiederholt, JL.; Bryant, BR. Gray oral reading tests. 4. Austin, TX: Pro-Ed; 2001.

- Wolf M, Katzir-Cohen T. Reading fluency and its intervention. *Scientific Studies of Reading* 2001;5:211–239.
- Woodcock, RW. Woodcock reading mastery test—Revised. Circle Pines, MN: AGS Publishing; 1998.

Table 1

Researcher-Identified and Teacher-Nominated Word Callers Across Studies 1 and 2 (N = 1,070)

Grade and criterion	<i>n</i>	%
Study 1 (<i>n</i> = 868)		
Grade 2 (<i>n</i> = 521)		
Criterion 1	4	0.8
Criterion 2	2	0.4
Grade 3 (<i>n</i> = 347)		
Criteria 1	8	2.3
Criteria 2	5	1.4
Study 2 (<i>n</i> = 202)		
Grade 3 (<i>n</i> = 110)		
Criterion 1	2	1.8
Criterion 2	0	0
Teacher nominated	25	21.7
Grade 5 (<i>n</i> = 92)		
Criterion 1	9	9.8
Criterion 2	1	1.1
Teacher nominated	20	22.7

Table 2

Correlation Between Reading Fluency and Reading Comprehension Across Studies 1 and 2 (N = 1,070)

Sample	<i>n</i>	<i>r</i>
Study 1		
Combined sample	868	.66**
Grade 2	521	.72**
Grade 3	347	.59**
Study 2		
Combined sample	202	.58**
Grade 3	101	.65**
Grade 5	92	.51**

**
 $p < .01$.

Table 3

Contingency Table for Teacher-Nominated Versus Researcher-Identified Word Callers in Study 2 (n = 202)

Teacher nomination	Researcher identified	
	Word caller	Nonword caller
Word caller	True positive = 3	False positive = 42
Nonword caller	False negative = 8	True negative = 149

Note. Criterion 1 was used by the researcher to identify word callers.

Table 4

Means and Standard Deviations for Teacher-Nominated Word Callers on Standardized and Teacher Assessments for Study 2 (n = 202)

Assessment	Word caller		Peer	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Standardized assessments				
Reading fluency (GORT-4)	90.65	15.98	97.0	18.26
Reading comprehension (WIAT-RC)	92.44	11.39	100.76	12.82
Teacher assessments				
Reading fluency (NAEP)	2.69	0.76	3.23	0.91
Reading comprehension (Scale)	1.64	0.61	2.40	0.70

Note. GORT-4 = Gray Oral Reading Test—Fourth Edition; WIAT-RC = Wechsler Individual Achievement Test—Reading Comprehension subtest; NAEP = National Assessment of Education Progress; Scale = Comprehension Scale.

Table 5

Contingency Table for Poor Readers Versus Teacher-Nominated Word Callers in Study 2 (n = 202)

Teacher nomination	Poor reader	Adequate reader
Word caller	11	34
Nonword caller	18	139