Canadian Journal of School Psychology
27(2) 150–165

DOI: 10.1177/0829573512438635
http://cjs.sagepub.com
SAGE

Phonological Segmentation Assessment Is Not Enough: A Comparison of Three Phonological Awareness Tests With First and Second Graders

David A. Kilpatrick[1]

Abstract

Despite extensive research on phonological awareness and reading, there has been little effort to study practical questions that would assist practitioners regarding the choice and interpretation of the phonological awareness tests available to them. This study examined the relationship between decoding (real and pseudowords) and three phonological awareness tests (segmentation, blending, and manipulation) taken from the *Comprehensive Test of Phonological Processing (CTOPP)* with an unselected population of first grade ($n = 67$) and second grade ($n = 49$) students. Segmentation displayed the weakest correlation with reading and accounted for no statistical variance in reading beyond what was found in the blending test. It also failed to account for a substantial amount of variance in reading that is captured by the manipulation test. Despite its popularity in educational contexts, phonological segmentation may be less useful than phonological manipulation or blending in assessing the phonological substrates of reading at these grade levels.

Resumé

Malgré des recherches approfondies sur la conscience phonologique et la lecture, peu d'efforts ont porté sur les questions pratiques qui pourraient guider les intervenants dans le choix et l'interprétation des tests de conscience phonologique qui leur sont

[1]State University of New York, College at Cortland, Cortland, NY, USA

Corresponding Author:
David A. Kilpatrick, State University of New York, College at Cortland, Department of Psychology, SUNY Cortland, P.O. Box 2000, Cortland, NY 13045, USA
Email: kilpatrickd@cortland.edu

offerts. Cette étude examine la relation entre le décodage (mots et pseudomots) et trois tests de conscience phonologique (segmentation, fusion et manipulation) tirés du Comprehensive Test of Phonological Processing (CTOPP) chez des élèves de première ($n = 67$) et de deuxième année ($n = 49$) du primaire. La segmentation affiche la corrélation la plus faible avec la lecture et n'apporte aucune contribution significative au modèle de régression linéaire au-delà de celle associée à la mesure de fusion. Elle n'arrive pas non plus à rendre compte de la variance en lecture contrairement au test de manipulation. Malgré la popularité de cet indicateur dans le milieu scolaire, la segmentation phonologique pourrait être moins utile que la manipulation ou la fusion phonologiques dans l'évaluation des substrats phonologiques de la lecture au primaire.

Keywords

phonological awareness assessment, phonological segmentation, phonological blending, phonological manipulation, Comprehensive Test of Phonological Processing (CTOPP)

Based on extensive evidence, researchers have determined that phonological awareness is strongly associated with the development of word-level reading skills (Perfetti, Beck, Bell, & Hughes, 1987; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Wagner, Torgesen, & Rashotte, 1994). Phonological awareness includes the ability to notice that spoken words can be divided into smaller units such as syllables, onsets, rimes, and phonemes. Students who develop phonological awareness to the phoneme level are able to quickly and easily map printed words to permanent memory (Høien, Lundberg, Stanovich, & Bjaalid, 1995; Laing & Hulme, 1999), while those who do not typically struggle in reading (Bruck, 1992; Greenberg, Ehri, & Perin, 1997; Vellutino et al., 2004).

The goal of the present article is to address a practical question that has been rarely addressed in the research literature. Simply put, Which phonological awareness test or tests will be most helpful in determining the presence of phonological awareness difficulties in educational contexts? Some practitioners may assume, based on the popularity of phonological segmentation tasks (e.g., DIBELS, AIMSweb, PALS, Yopp-Singer), that such popularity stems from a body of best practice research. This is not the case. The construct of phonological awareness has been evaluated in multiple ways, such as segmentation, blending, categorization, and manipulation. It has yet to be established whether one phonological awareness test or task is better than another at determining if a student's reading progress is being affected by poor phonological awareness skills.

While best practice cannot be established by a single study, the goal here is to take an important step toward raising this issue as well as encouraging further inquiry. Answering this best practice question is difficult for two reasons: (a) the sheer number

of tasks that have been used to assess the construct of phonological awareness, and (b) the lack of research that directly addresses this question.

Measuring the Construct of Phonological Awareness

Researchers have measured the construct of phonological awareness in many ways, such as rhyming, segmentation, blending, isolation, categorization, and manipulation (Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003; Chafouleas, Lewandowski, Smith, & Blachman, 1997; Høien et al., 1995; Lenchner, Gerber, & Routh, 1990; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Schatschneider, Francis, Foorman, Fletcher, Mehta, 1999; Stanovich, Cunningham, & Cramer, 1984; Vloedgraven, & Verhoeven, 2009; Yopp, 1988). In addition, each of these ways of measuring of phonological awareness has been examined using multiple tasks. For example, rhyming can involve rhyme *recognition*, rhyme *matching*, or rhyme *production* (Stanovich et al., 1984; Yopp, 1988) while manipulation can involve *deletion*, *substitution*, or *reversal* (Lenchner et al., 1990; Wagoner et al., 1999). Also, across the various tasks, other factors have been considered such as levels of linguistic complexity (syllables, onset-rimes, & phonemes), position of phonemes within words, whether a phoneme is part of a blend, or whether it is voiced or unvoiced (Anthony et al., 2003; Seymour & Evans, 1994; Stahl & Murray, 1994). These factors affect performance on phonological awareness tasks, making best practice questions rather complex.

The Paucity of "Best Practice" Research

Despite hundreds of studies on the relationship between phonological awareness and reading, there has been no concerted effort devoted to determining the most practical and effective way of evaluating phonological awareness in schools, given the assessment instruments available to educational professionals. Numerous studies have incorporated multiple phonological awareness tasks (e.g., Anthony et al., 2003; Høien et al., 1995; Schatschneider et al., 1999, 2004; Seymour & Evans, 1994; Stahl & Murray, 1994; Vloedgraven, & Verhoeven, 2009; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner et al., 1994; Yopp, 1988). However, these studies made no attempt at directly comparing tasks for clinical utility. Rather they used multiple measures to either determine the factor structure of phonological awareness, or to create a phonological awareness factor that is then used to study its relationship with reading. Two studies, however, are welcome exceptions, Chafouleas et al. (1997) and Swank & Catts (1994). In 1994, Swank and Catts could say, "it remains unclear which measures of phonological awareness will be the most effective in clinical practice for identifying children who lack sufficient phonological awareness" (p. 10). This comment is as relevant today as it was in 1994. Unfortunately, these two studies have not had a substantial impact on the field, based on how rarely they have been cited by later researchers (according to the citations forward feature in PsychINFO).

The National Early Literacy Panel (2008) conducted a meta-analysis of studies looking at the relationship between reading and different aspects of phonological awareness. This is a welcome step because they raised the question of differing strengths of correlation between reading and differing aspects of the construct of phonological awareness. However, they examined the correlations between reading and two *dimensions* of the phonological awareness construct (level of linguistic complexity & analysis vs. synthesis) abstracted from various tasks, rather than a direct comparison between the tasks themselves. Thus, the NELP report does not provide assistance in terms of selecting among the various tests available to practitioners.

The Context for the Study

A phonological awareness test can be used as a screening (e.g., DIBELS). But once reading skills begin to develop, direct reading-related tasks (word identification, phonic recoding) tend to parallel or eclipse phonological awareness as a predictor of future reading (NELP, 2008). Phonological awareness tests are also administered to students referred for an evaluation of reading difficulties. In such instances, they are not used to predict future reading skills but to determine the likelihood that the student's level of phonological awareness development is affecting his or her reading progress. The question in this context is whether some phonological awareness tests or tasks (e.g., segmentation, blending, manipulation) are more closely associated with early reading skills than others, and therefore presumably better at addressing the question that prompted the assessment.

A related question is how to interpret the profile of a battery of phonological awareness tests, such the CTOPP. *There is simply no specific research literature designed to assist with this practical issue*. If a child's performance on the phonological manipulation, blending, and segmentation subtests are all consistently low or high, one may feel confident in deciding whether phonological awareness training is needed. However, it is quite common for a student to display a mixed profile with these subtests. What can be concluded from such a profile? Should teachers invest valuable instructional time in phonological awareness training with a student who displays a mixed profile, or only provide intervention for those with consistently low phonological awareness scores? Can such decisions be data driven? At the present time, they cannot, because other than the studies mentioned above (Chafouleas et al., 1997; Swank & Catts, 1994), there has been no effort to investigate this specific question. It would seem that a direct comparison between these phonological awareness tests and reading might begin to address this issue. The present study was designed to make such a direct comparison.

The impetus for this article was the author's experience of administering the CTOPP to hundreds of students referred for reading difficulties in an elementary school context. It became clear that approximately half of these students performed at or above the 50th percentile on the segmentation test, while an estimated 80% of these same students performed low average to below average on the manipulation task

(which involves deleting sounds from words). This suggested that these two types of phonological awareness tests were not equally well suited for determining whether a student's reading difficulty was the result of weak phonological awareness skills. Phonological segmentation assessment is the lone measure of phonological awareness within several popular tests and batteries (e.g., DIBELS, AIMSweb, PALS, Yopp-Singer, Sawyer STAS). The concern is that if weak readers do well on a segmentation test, it may be assumed poor phonological awareness is not involved, while a test of phonological manipulation might have suggested otherwise. On the other hand, it may be that segmentation tests more accurately reflect the phonological substrates of reading acquisition but that manipulation tasks overidentify phonological awareness difficulties. A better understanding of this practical issue prompted the literature review summarized below and the present empirical study.

Addressing the Issue of Best Practice

This article is designed to take an initial step toward examining which test(s) might be the most clinically useful in determining if children have phonological awareness difficulties. It uses a subset of the many phonological awareness assessment approaches described above. It therefore represents an early step in addressing the best practice questions raised here. Also, most previous research on various phonological awareness tasks examined just that: phonological *tasks*. Many, if not most, of the tasks used were researcher-designed and do not represent actual tests available to school-based evaluators. Even the two studies referred to above that attempted to address clinical usefulness (Chafouleas et al., 1997; Swank & Catts, 1994) used experimenter-designed tasks that predate the CTOPP and are not commercially available in the form used in those studies. *Practitioners need information regarding the relative usefulness of actual tests available to them, an issue that has received virtually no attention within the extensive phonological awareness literature*. Therefore, the present study looks at three subtests from the commercially available *Comprehensive Test of Phonological Processing (CTOPP*; Wagner, Torgesen, & Rashotte, 1999). Each subtest represents a different phonological awareness task (segmentation, blending, and manipulation).

Segmentation, blending, and manipulation were selected based on empirical and practical considerations. Rhyming was excluded because studies have shown it has little or no discriminant validity beyond kindergarten and all participants in this study were beyond kindergarten. Phoneme categorization and phoneme isolation are not as commonly found in tests available for educators. Also, Oakhill and Kyle (2000) found that phonological categorization is confounded with working memory, but phonological manipulation is not. This leaves segmentation, blending, and manipulation.

Manipulation versus segmentation. While little attention has previously been drawn to this fact, numerous research reports include data to show that from first grade and beyond, manipulation tasks display higher correlations with reading measures than

segmentation tasks (Backman, 1983; Kroese, Hynd, Knight, Hiemenz, & Hall, 2000; Lenchner et al., 1990; Perfetti et al., 1987; Swank & Catts, 1994; Wagner et al., 1993). Authors rarely mention this difference. One must discover these differences by examining reported correlation tables. Rare exceptions include Catts, Fey, Zhang, & Tomblin (2001) who said phonological manipulation "ranks highly among phonological awareness tasks in predicting reading achievement" (p. 40) and Lenchner et al. (1990) who stated that their manipulation task had a higher correlation with decoding ($r = .78$ and $r = .74$) than any segmentation task reported in the literature.

The following study compares three tests representing segmentation, blending, and manipulation with the reading abilities of first and second graders. It is in these grades that reading difficulties are commonly discovered, so they seemed appropriate grade levels to consider. An unselected sample was used rather than a clinical sample to provide an examination of how these phonological awareness tests correlate with the continuum of reading skill levels among first and second graders. A clinical sample might skew the correlations and be more appropriate for addressing different aspects of the question of clinical utility than are being addressed here. The goal in this article is to take a first step toward addressing best practice questions by seeking explicit confirmation of what is implicitly reported elsewhere, which is that different phonological awareness tests more closely parallel the development of early reading skills than others.

Based on the correlations reported in previous studies, it is predicted that at both grade levels, the segmentation test will have lower correlations with word-level reading and provide minimal help in clinical assessment beyond what can be gleaned from tests of manipulation and blending.

Method

Participants

Participants were 67 first-grade (30 female, 37 male) and 49 second-grade (23 female, 26 male) students from a lower-middle class suburban elementary school in Upstate New York. No specific data was collected on race, but districtwide, more than 94% of the students are White. All students in first and second grade were recruited and there were no preselection criteria other than the absence of any visual, hearing, or cognitive disabilities. All students were native speakers of English. Five of the first graders and eight of the second graders had already been identified as having a reading disability.

Materials

Reading tests. All participants received the *Word Identification* and *Word Attack* subtests from the *Woodcock Reading Mastery Test–Revised* (WRMT-R; Woodcock, 1999). In the *Word Identification* subtest, students are asked to read a graded word list.

The *Word Attack* subtest involves reading pseudowords (e.g., *seeg, trast*) of increasing difficulty. The participant's scores on these tests were the total number of items read correctly.

Phonological awareness tests. Three phonological awareness tests were administered. These tests were taken from the *CTOPP: Segmenting Words*, *Blending Words*, and *Elision*, which evaluate phonological segmentation, blending, and manipulation, respectively. *Segmenting Words* involves separating words into their individual sounds (e.g., "Say *sat* one part at a time" = /s/ /æ/ /t/). *Blending Words* involves identifying a word from its parts (e.g., "What word do these sounds make: /t/ /æ/ /n/?" = *tan*). The *CTOPP Elision* subtest involves deleting a sound from a word (e.g., "Say *drive* without the /r/" = *dive*). Both *Elision* and *Blending Words* take students through the continuum of linguistic complexity, starting with syllable items, progressing to onset-rime items and phoneme-level items. *Segmenting Words* uses only phoneme-level items, but begins with two phoneme words and progresses to words with more phonemes.

Procedure

The *CTOPP* measures were administered first, in standard *CTOPP* order (*Elision*, *Blending Words*, *Segmenting Words*). Because the goal was to determine the practical usefulness of these tests, preserving the order that evaluators would actually use was deemed essential. The *WRMT-R* subtests were administered afterward. Students were pulled from independent work time and tested in a hallway outside their classrooms. Each session lasted about 15 min. All data were gathered from December to March by a certified school psychologist.

Results

Table 1 includes the means and standard deviations for the raw scores and standard scores for each of the reading and phonological awareness tests, at both grade levels. Raw scores were used in all analyses but standard scores are also reported in Table 1 to provide a normative comparison of these unselected samples. Table 2 presents the intercorrelations among the measures. All phonological awareness tests correlated significantly with both Word Identification and Word Attack at both grade levels. At both grade levels, Segmenting Words had the weakest correlation with both reading measures. Also at both grade levels, the three phonological awareness subtests were significantly intercorrelated, except for Elision and Segmenting Words at second grade.

To explore the relationship among these measures, hierarchical multiple regression analyses were conducted separately on the first and second grade samples (Table 3), one analysis using *Word Identification* as the dependent variable, and the other using *Word Attack*. At both grade levels, *Segmenting Words* accounted for no unique variance beyond *Blending Words* (Model 2). It did contribute unique variance with first graders when the model included *Segmenting Words* and *Elision* (Model 3), though it

Table 1. Means and Standard Deviations: Grades 1 and 2

	Grade 1 ($n = 67$)		Grade 2 ($n = 49$)	
Measure	Mean	(*SD*)	Mean	(*SD*)
Raw scores				
Age	6 years 6 months	(3.7 m)	7 y10m	(3.7 m)
WRMT-R Word Identification	28.18	(15.84)	50.88	(16.37)
WRMT-R Word Attack	12.31	(8.00)	24.77	(9.15)
CTOPP Elision	7.69	(3.32)	10.86	(4.68)
CTOPP Segmentation	8.12	(2.79)	8.18	(3.19)
CTOPP Blending Words	12.97	(3.56)	14.24	(2.93)
Standard scores/scaled scores				
WRMT-R Word Identification	114.43	(12.95)	106.87	(12.34)
WRMT-R Word Attack	113.29	(9.91)	112.51	(14.89)
CTOPP Elision	11.21	(2.22)	10.28	(2.68)
CTOPP Segmentation	9.46	(1.41)	9.53	(1.57)
CTOPP Blending Words	13.27	(2.53)	11.68	(1.98)

Note: (1) WRMT-R = Woodcock Reading Mastery Test–Revised; CTOPP = Comprehensive Test of Phonological Processing. (2) The standard scores for the first graders on the WRMT-R subtests likely reflect an inflated representation of actual skills. Participants in this study are from New York State, which has a later cut-off for kindergarten entry than most other states (age 5 by December 1; September 1 is most common). Therefore, the national norms of the WRMT-R compare these students with students who, on average, have completed 3 to 4 fewer months of schooling.

Table 2. Subtest Intercorrelations: Grade 1 ($n = 67$) and Grade 2 ($n = 49$)

	WID	WA	EL	SEG	BW
WID		.88***	.56***	.31*	.64***
WA	.80***		.67***	.33*	.51***
EL	.60***	.59***		.20	.29*
SEG	.47***	.42***	.26*		.35*
BW	.65***	.57***	.47***	.55***	

Note: Grade 1 is below the diagonal; Grade 2 is above. WID = WRMT-R Word Identification; WA = WRMT-R Word Attack; EL = CTOPP *Elision*; SEG = CTOPP *Segmenting Words*; BW = CTOPP *Blending Words*.
*$p < .05$. **$p < .01$. ***$p < .001$.

just failed to reach significance at second grade. In the model that included all three tests, an impressive 55% of the variance in reading was accounted for by these phonological awareness measures. *Blending Words* and *Elision* each contributed a substantial amount of unique variance in this model, though *Segmenting Words* does not. Using *Word Attack* as the dependent variable, the pattern of results was nearly identical, with only slight variations in magnitude.

Table 3. Regression Analyses: Grade 1 ($n = 67$) and Grade 2 ($n = 49$)

Dependent Variable	Model	Independent variables	Grade 1 β	Grade 1 *p*	Grade 1 R^2	Grade 2 β	Grade 2 *p*	Grade 2 R^2
WRMT-R Word Identification								
	1	CTOPP Segmenting Words	.47	<.001	.22	.31	.03	.097
	2	CTOPP Segmenting Words	.17	.15 (*ns*)		.11	.37 (*ns*)	
		CTOPP Blending Words	.56	<.001	.44	.57	<.001	.38
	3	CTOPP Segmenting Words	.34	.001		.21	.09 (ns)	
		CTOPP Elision	.51	<.001	.47	.51	<.001	.35
	4	CTOPP Segmenting Words	.17	.10 (*ns*)		.06	.58 (*ns*)	
		CTOPP Blending Words	.37	.002		.48	<.001	
		CTOPP Elision	.38	<.001	.55	.42	<.001	.54
WRMT-R Word Attack								
	1	CTOPP Segmenting Words	.44	<.001	.20	.33	.02	
	2	CTOPP Segmenting Words	.17	.16 (*ns*)		.19	.18 (*ns*)	
		CTOPP Blending Words	.50	<.001	.37	.42	.004	.26
	3	CTOPP Segmenting Words	.31	.002		.21	.06 (*ns*)	
		CTOPP Elision	.53	<.001	.46	.62	<.001	.48
	4	CTOPP Segmenting Words	.18	.10 (*ns*)		.12	.28 (*ns*)	
		CTOPP Blending Words	.29	.02		.29	.01	
		CTOPP Elision	.42	<.001	.50	.56	<.001	.55

Note: WRMT-R = Woodcock Reading Mastery Test–Revised; CTOPP = Comprehensive Test of Phonological Processing.

Discussion

At both grade levels, *Segmenting Words* displayed lower correlations with both word identification and phonic decoding tasks. Also, it contributed no unique variance to these reading tests once its overlap in variance with *Blending Words* has been partialled out. With first graders, and with a nonsignificant tendency in second graders, *Segmenting Words* and *Elision* each account for unique variance when the model includes just those two (Model 3). This finding seems to suggest that because *Elision* was more highly correlated with reading but contributed much variance beyond *Segmenting Words*, *Segmenting Words* failed to capture a large portion of variance that could be attributed to the construct of phonological awareness. To a lesser extent (based on Beta scores), something similar could be said about *Elision*, at least with the first grade sample. *Elision* failed to account for a portion of variance in reading ability captured by *Segmenting Words*. This suggests that neither test can stand alone, particularly at first grade. *Elision* and *Blending Words* each account for unique variance in word-level reading tests (Model 4), and together account for an impressive

amount of variance in both real word and pseudoword decoding. This also suggests that for the population of students in the present study, none of the three phonological awareness subtests stands alone but that *Segmenting Words* accounts for no variance in word-level reading beyond what is found in the other tests.

These results suggest that *Segmenting Words* may be the least helpful in determining whether a student's word-level reading difficulties stem from phonological awareness deficits. It accounts for no unique variance beyond *Blending Words* and fails to account for a large amount of variance in word reading captured by *Elision*, which is a test that is much more highly correlated with reading.

To understand the differences in findings regarding *Segmenting Words* between the first- and second-grade samples, it may be instructive to note that the raw scores of the *Segmenting Words* subtest were nearly identical between the first- and second-grade samples (see Table 1). This implies that phonological segmentation skills plateau between the first and second grades, a finding reported by others (e.g., Vloedgraven & Verhoeven, 2009; Wagner et al., 1993). By contrast, the mean raw scores of the second graders on the *Elision* and *Blending Words* subtests were higher than the first-grade sample. This appears to indicate that phonological awareness skills continue to develop beyond first grade but that a task like *Segmenting Words* loses some of its strength of correlation with reading growth and development after first grade.

Implications

The data presented above suggests that while phonological segmentation is commonly incorporated into popular test batteries, it may not be best practice to use it alone to determine whether a student may have difficulties with phonological awareness. Both the CTOPP *Elision* and *Blending Words* subtests, and particularly the combination of the two, appear to be superior in accessing the phonological substrates of reading than *Segmenting Words*. Though other studies may not have made the direct comparisons this study made, the present findings are similar to what has been reported in the correlation tables of previous reports (e.g., Backman, 1983; Kroese et al., 2000; Lenchner et al., 1990; Perfetti et al., 1987; Wagner et al., 1993).

From this it might be reasonable to assume that if a weak reader does poorly on a manipulation task or a blending task but does well on a segmentation task that the student is likely to have phonological awareness difficulties. In such a case, the segmentation task is simply not as helpful in detecting these difficulties. Scatter plots from the first- (Figure 1) and second-grade (Figure 2) samples may help illustrate this issue. Figure 1a shows that there are many first graders who are among the lower readers in this sample who are at or above the group's median in their segmentation skills. By contrast, no children from among the lower first-grade readers performed at or above their group's median on the *Elision* subtest (Figure 1c). This confirms the clinical observations that prompted this investigation. However, there were numerous students who appear to be doing well in reading that had a comparatively weak score on the *Elision* subtest,[1] while no such pattern emerged with either *Segmenting Words* or *Blending Words*.

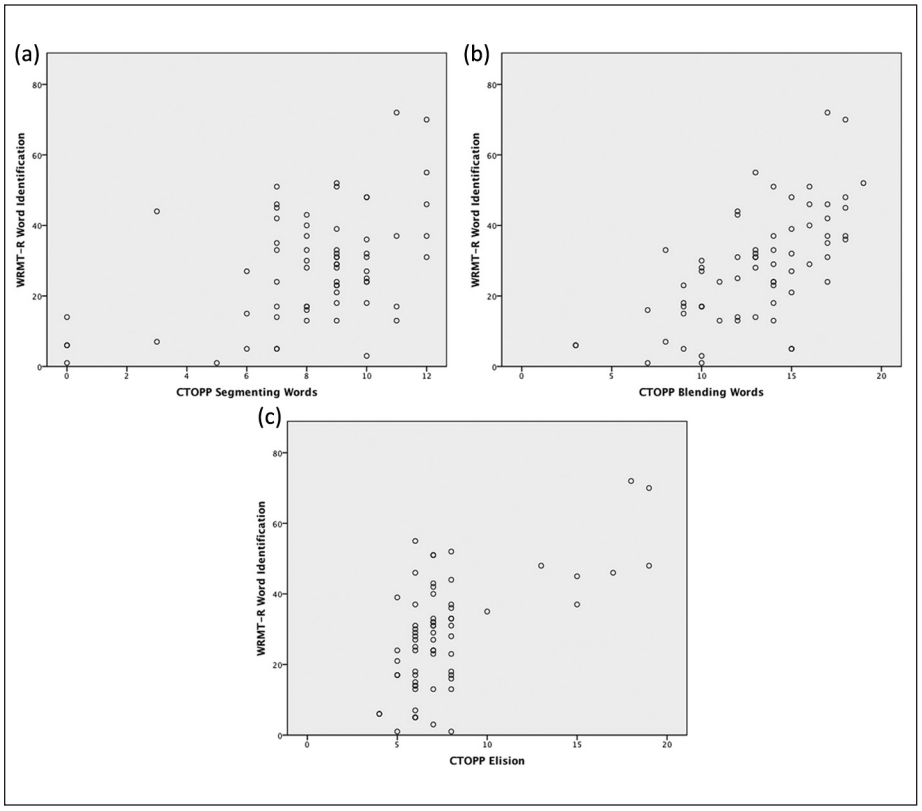

Figure 1. Grade 1 scatter plots (n = 67)
Note: WRMT-R = Woodcock Reading Mastery Test–Revised; CTOPP = Comprehensive Test of Phonological Processing.

One implication of the findings above is that a segmentation task may fail to recognize phonological awareness difficulties in a struggling reader. Figure 1b indicates that *Blending Words* also appears to display a similar pattern as *Segmenting Words*, despite its stronger correlation with reading. It must be noticed that for *Blending Words*, and to a slightly lesser extent *Segmenting Words*, students who display a relatively weak performance on these tasks are almost invariably weaker readers. The clinical implication of this is that weak performance on either of these subtests seems indicative of genuine phonological awareness difficulties, while a weak performance on *Elision* might not. However, good performance on either *Segmenting Words* or *Blending Words* cannot be relied on to accurately rule out phonological awareness difficulties.

Educators are under pressure to use "research-based" approaches when assessing and teaching reading. Nearly all phonological awareness tasks can be called "research based" because they all have been shown via research to correlate with reading at some age level or another. The question, however, is about best practice.

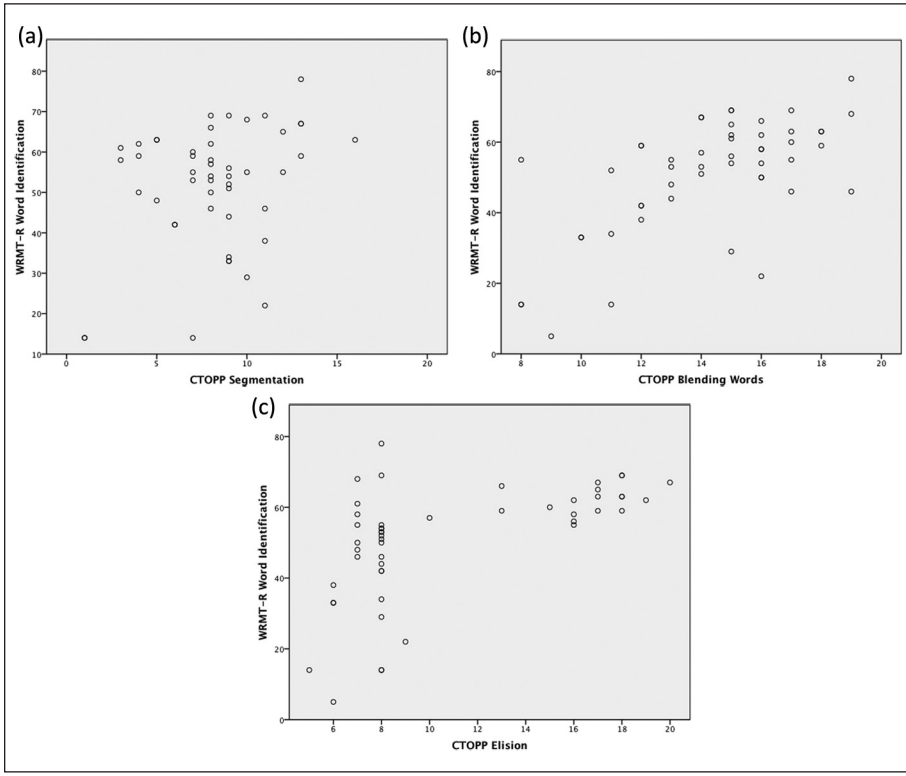

Figure 2. Grade 2 scatter plots (n = 49)
Note: WRMT-R = Woodcock Reading Mastery Test–Revised; CTOPP = Comprehensive Test of Phonological Processing.

Among weak readers, an important clinical question that evaluators seek to answer is whether those reading difficulties are affected by deficits in phonological awareness. Such deficits in phonological awareness can hinder reading skills throughout adolescence (Lenchner et al., 1990) and into adulthood (Bruck, 1992; Greenberg et al., 1997). So when students present themselves with reading difficulties, it is imperative that practitioners use a test most likely to determine if phonological awareness is a factor contributing to the reading difficulties. *This means that if educators rely on segmentation or blending to assess phonological awareness in reading assessments, they may fail to recognize phonological awareness difficulties in a meaningful percentage of students with such difficulties*. If phonological awareness difficulties are "ruled out" based on average performance on a segmentation or blending task, students are unlikely to get the phonological awareness training they need to assist them in their reading progress.

While these data need further confirmation, the implications are that educators should reexamine the common practice of relying on segmentation assessment. Rather,

a combination of manipulation and blending tasks is likely to provide a better assessment of a student's phonological awareness skills. Each of these tasks captures unique variance in reading skill and together they account for a sizeable amount of that variance. It would seem advisable that if subtest performance is low on both manipulation and blending tests, then phonological awareness intervention is indicated. If both are average or better, it may be reasonable to assume the student's reading difficulties are unrelated to phonological awareness. The difficulty comes when the performance is split. Here, further investigation will be needed to explore this question more fully.

This study has its limitations. Only the first- and second-grade levels were evaluated. It would be ideal to investigate preschoolers through third graders, as well as older, struggling students. Also, this study used concurrent measures of phonological awareness and reading while longitudinal studies would also help address "best-practice" questions. However, in terms of strength of correlation, longitudinal studies are consistent with the present results (Wagner et al., 1994; Swank & Catts, 1994). Another important follow-up would be to compare a clinical sample with a typical sample and evaluate classification accuracy of the various phonological awareness tests.

A potential threat to the validity of these results was the lack of counterbalancing, which is designed to guard against practice effects or any other effect that a given order of administration may produce. However, three factors suggest that the order of administration did not likely have an impact on the present results. First, this study was ultimately an examination of specific phonological tests rather than more generically the tasks that make up those specific tests (i.e., segmentation, blending, manipulation). As an applied study, preserving the order of administration of the *CTOPP* was essential because practitioners would use that order in clinical practice. It could thus be argued that counterbalancing might have threatened the validity of the results. Second, the correlations between the specific phonological awareness tests and reading found in this study are consistent with findings from other studies. Third, the empirical findings suggest there was no practice effect. Rather, there was a decreasing correlation with reading across the three subtests administered. *Segmenting Words* was the third test administered and it consistently displayed weaker correlations with reading than the two prior tests. The only possibility here is that some sort of "reverse practice effect" occurred. The consistency of our results with previous studies into these types of tasks suggests that this is not likely.

Despite its limitations, the present article was designed to do two things. The first was to bring to the attention of researchers and practitioners what has already been reported but unheralded in existing research: From first grade and beyond, phonological segmentation tasks have weaker correlations with word-level reading skills than phonological manipulation tasks. Second, the present study provides explicit confirmation of what has already been previously reported implicitly. Educators need recommendations regarding which test best captures the construct of phonological awareness. It is therefore important to direct research attention explicitly at comparisons between differing phonological awareness tasks and tests to determine which one(s) would be most practical and most highly recommended for educators. While

the present study fell short of establishing best practice, it took a step in that direction by suggesting what is *not* likely to be best practice, which is relying exclusively on phonological segmentation, despite its popularity in the schools. This is not to say the skill of phonological segmentation is unimportant for reading, because it is *very* important. Rather, the point is manipulation and blending tasks appear to do a better job of accessing the construct of phonological awareness than a simple segmentation task.

It may be too obvious to state that we do not need another study showing that phonological awareness correlates with (or predicts) word-level reading skills. However, we *do* need more research on the relationship between phonological awareness and reading that addresses questions of clinical utility. The goal of this article was to (a) alert practitioners that it would be inadvisable based on current evidence to rely on phonological segmentation to assess the construct of phonological awareness beyond kindergarten; and (b) serve as a catalyst for researchers to address the best practice questions. It will take numerous studies to develop an empirical base to address the best practice questions. This study was designed to take an important step in that direction.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Note

1. The pattern displayed with the *Elision* subtest might be an artifact of its administration and not reflective of the degree to which phonological manipulation approximates the construct of phonological awareness. The *Elision* subtest unfortunately includes a significant shift in task demands at Item 9, but feedback on incorrect responses stops after Item 5. Item 9 is the first to include phoneme deletion from within the middle of a two-syllable word (e.g., "Say *tiger* without saying /g/"). There are no instructions or sample items for this type of manipulation. The next two items are similar. Because the ceiling is reached following three incorrect items in a row, it is common for students to reach the ceiling between Items 9 and 11. When considering the pattern in Figures 1c and 2c, it is quite possible the "spikes" in the upper left are a reflection of this. It may be that some students simply were not clear about the sudden change in task demands. By contrast there are no "spike" patterns in the scatter plots of the other two subtests nor are there any shifts in task demands. This might mean that *Elision* has the potential of being an even more powerful test of phonological awareness were it not for this apparent artifact of test administration.

References

Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly, 38*, 470-487.

Backman, J. (1983). The role of psycholinguistic skills in reading acquisition: A look at early readers. *Reading Research Quarterly, 18*, 466-479.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*, 874-886.

Catts, H. W., Fey, M. E., Zhang, X., Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implication. *Language, Speech, and Hearing Services in Schools, 32*, 38-50.

Chafouleas, S. M., Lewandowski, L. J., Smith, C. R., & Blachman, B. A. (1997). Phonological awareness skills in children: Examining performance across tasks and ages. *Journal of Psychoeducational Assessment, 19*, 216-226.

Greenberg, D., Ehri, L. C., & Perin, D. (1997). Are word-reading processes the same or different in adult literacy students and third-fifth graders matched for reading level? *Journal of Educational Psychology, 89*, 262-275.

Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I.-K. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal, 7*, 171-188.

Kroese, J. M., Hynd, G. E., Knight, D. F., Hiemenz, J. R., & Hall, J. (2000). Clinical appraisal of spelling ability and its relationship to phonemic awareness (blending, segmenting, elision, and reversal), phonological memory, and reading in reading disabled, ADHD, and normal children. *Reading & Writing: An Interdisciplinary Journal, 13*, 105-131.

Laing, E., & Hulme, C. (1999). Phonological and semantic processes influence beginning readers' ability to learn to read words. *Journal of Experimental Child Psychology, 73*, 183-207.

Lenchner, O., Gerber, M. M., & Routh, D. K. (1990). Phonological awareness tasks as predictors of decoding ability: Beyond segmentation. *Journal of Learning Disabilities, 23*, 240-247.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

Oakhill, J., & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology, 75*, 152-164.

Perfetti, C. A., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283-319.

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*, 439-449.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*, 265-282.

Seymour, P. H. K., & Evans, H. M. (1994). Levels of phonological awareness and learning to read. *Reading and Writing: An Interdisciplinary Journal, 6*, 221-250.

Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38*, 175-190.

Swank, L. K., & Catts, H. W. (1994). Phonological awareness and written word decoding. *Language, Speech, and Hearing Services in Schools, 25*, 9-14.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2-40.

Vloedgraven, J. M. T., & Verhoeven, L. (2009). The nature of phonological awareness throughout the elementary grades: An item response theory perspective. *Learning and Individual Differences, 19*, 161-169.

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85*(1), 83-103.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*(1), 73-87.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *The Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: Pro-ED.

Woodcock, R. W. (1999). *Woodcock Reading Mastery Test–Revised*. Circle Pines, MN: American Guidance Services.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*, 159-177.

Bio

David A. Kilpatrick, PhD, is an assistant professor of psychology for the State University of New York, College at Cortland. In addition, he is a part-time school psychologist with the East Syracuse-Minoa School District in Upstate New York. His research interests include the development of word recognition skills and the phonological processes that underlie reading disabilities. He received his doctorate from Syracuse University.