The Phonological Awareness Screening Test (PAST): An Initial Report

David A. Kilpatrick

State University of New York, College at Cortland

East Syracuse-Minoa Central School District

Philip J. McInnis

Keuka College & Nazareth College

Hammondsport Central School District

Abstract

The *Phonological Awareness Screening Test* (PAST) is here presented as a valid and reliable assessment of phonological awareness that takes 5-8 minutes to administer. The data presented demonstrates that the PAST tends to correlate more strongly with word-level reading than currently available phonological awareness tests. The PAST also displays strong correlations with other tests of phonological awareness. It is offered to educational professionals and researchers as a free, public domain test. The PAST represents the third generation of the classic Rosner & Simon (1971) *Auditory Analysis Test* (AAT), but revised based on research over the last three decades as well as extensive clinical use in school settings. A review of research on phonological awareness assessment is followed by data that suggests that the PAST is a valid and reliable instrument based upon samples of typical students from kindergarten, first, second, and fifth graders, and college students along with samples of referred students from first to eighth grade. Finally, instructions for administration, scoring, and interpretation are offered.

The Phonological Awareness Screening Test (PAST): An Initial Report

> "The ability to differentiate the sounds of the language, both in listening and speaking, has long been recognized as an important factor in learning to read. The literature concerned with the teaching of reading has repeatedly acknowledged the importance of auditory perception."

So began an article in the *Journal of Learning Disabilities* 43 years ago by Jerome Rosner & Dorothea Simon (Rosner & Simon, 1971). The article was entitled "The Auditory Analysis Test: An Initial Report." The present article is designed not simply to emulate that report, but to provide a practical update of that "initial report," with a version of their Auditory Analysis Test (*AAT*) that has been in use, in one form or another, for nearly four decades.

Rosner and Simon's *AAT* has its modern derivatives, including the *Elision* subtest from the *Comprehensive Test of Phonological Processing* (*CTOPP*; Wagner, Torgesen, & Rashotte, 1999; second edition 2013), a syllable/phoneme deletion task developed by Catts (1993) and colleagues and used in various studies (e.g., Catts, Fey, Zhang, & Tomblin, 1999, 2001) as well as in other experimenter designed versions used in studies examining phonological awareness (e.g., Hulme, et al., 2002; Laing & Hulme, 1999).

Our goal here is to present data on the *Phonological Awareness Screening Test* (*PAST*), a assessment of phonological awareness that is a direct descendent of the *AAT*. The test is presented here, with preliminary data supporting its validity and reliability, along with instructions for administration. It will be shown below that the PAST offers educational evaluators and researchers a tool that typically correlates more strongly with word-level reading and phonics skills than most phonological awareness assessments in current use and can function as a valuable supplement to existing normed test batteries. Following the tradition of the AAT, the intent of this report is to make the PAST a public domain test for practitioners and researchers to use in the assessment of phonological awareness.

To better understand the PAST, the nature of phonological awareness and its importance in reading acquisition will be briefly reviewed. Next, we will provide a brief history of the PAST assessment, along with information about the type of phonological awareness task utilized in the PAST. After this, the empirical data that supports the validity of the PAST will be presented. Finally, instructions for scoring and interpretation are included. This is intended to assist educators and researchers to make use of the PAST both in school-based evaluations as well as in research studies.

The Importance of Phonological Awareness

Based upon 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. Phonemic awareness represents the most sophisticated form of phonological awareness, and the focus is on the phonemes within words. Phonemes are the smallest units of spoken language. In alphabetic languages, phonemes are typically represented by a single letter, though in English, phonemes are often represented by more than one letter (e.g., *ch–, sh–, –igh, –ck*). So, for example, most young children eventually notice that the word *red* has three phonemes (/r/ /e/ /d/) while the word *shoe* has two (/ʃ/ /u/).

Students who develop phonological awareness are able to quickly and easily map printed words to permanent memory (Cardoso-Martins, Mamede Resende, & Assunção Rodrigues, 2002; Dixon, Stuart, & Masterson, 2002; Ehri, 2005; Høien, Lundberg, Stanovich, & Bjaalid, 1995; Laing & Hulme, 1999). Those who do not develop phonological awareness typically struggle in reading (Greenberg, Ehri, & Perin, 1997; Vellutino, et al., 1996; Vellutino, et al., 2004).

Because phonological/phonemic awareness is critical for reading, it seems important for educators to test or screen for a student's phonological awareness skills (National Reading Panel, 2000). One difficulty educators face, however, is that is that there are many ways to assess phonological awareness.

Phonological awareness assessment tasks

Researchers have measured the construct of phonological awareness in a variety of ways, including rhyming, segmenting, blending, isolating, categorizing, and manipulating sounds in words (Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003; Chafouleas, Lewandowski, Smith, & Blachman, 1997; Høien, et al., 1995; Lenchner, Gerber, & Routh, 1990; Lundberg, Olofsson, & Wall, 1980; Stanovich, Cunningham, & Cramer, 1984; Yopp, 1988). *Rhyming,* as the name implies, expects children to perform a task involving rhyming or rhyme recognition. *Segmentation* consists of breaking a word into segments. With *blending* a student is given parts of words and he or she must indicate the word he would get if he "blended" those sounds together (e.g., the sounds /t/ /r/ /i/ /p/ form *trip*). *Isolation* involves determining the position of a sound within a word. For example, a student may be asked where the /d/ sound is in *dog* or *bed*. *Categorization* is most commonly represented by the *oddity task,* in which the student must indicate which word begins or ends with a different sound than the others (e.g., “Which word ends with a different sound than the others: *bike, brush, truck*?”).

By contrast to these ways of assessing phonological awareness, the *AAT* and the *PAST* use *phonological manipulation*. Phonological manipulation can involve deleting sounds from words, substituting sounds, or reversing sounds (Kroese, Hynd, Knight, Hiemenz, & Hall, 2000; Lundberg, et al., 1980; McInnis, 1981, 1999; Wagner, Torgesen, Rashotte, 1999). Deletion appears to have been the earliest reported form of phonological manipulation (Bruce, 1964; Rosner & Simon, 1971).

Judging from its inclusion in many phonological awareness batteries, phonological segmentation is arguably the most widely used phonological awareness assessment in public education (e.g., DIBELS, AIMSweb, PALS, easyCBM, Yopp-Singer). One may infer that the decision to use segmentation rather than one of the other ways of assessing phonological awareness (i.e., blending, isolation, manipulation) was based upon a body of best practice research. This is not the case. Despite the hundreds of research studies on the relationship between phonological awareness and reading, to date there has been no concerted effort to determine which of the many possible approaches to phonological awareness assessment is most clinically useful to educational professionals seeking to determine if phonological awareness plays a role in a student’s reading difficulties.

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, et al., 1993; Wagner, et al., 1994; Yopp, 1988). However, these studies made no attempt to directly compare tasks for clinical utility. Rather they used multiple measures of phonological awareness 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 development. In two of the rare instances in which specific tasks (e.g., segmentation, blending, manipulation) were examined for clinical utility (Kilpatrick, 2012; Swank & Catts, 1994), segmentation had lower correlations with word level reading than blending and manipulation among first and second graders, and contributed no unique variance in explaining reading beyond those other two tasks.

These explicit findings roughly parallel what has been found implicitly in the phonological awareness literature. While this fact has gone largely unheralded, 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, et al., 2000; Lenchner, et al., 1990; Perfetti, et al., 1987; Swank & Catts, 1994; Wagner, et al., 1993). Authors rarely mention this difference, so one must discover these differences by examining their correlation tables. Rare exceptions include Catts, et al. (2001) who explicitly stated 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 & r = .74) than any segmentation task reported in the literature.

An informal task analysis suggests why this may be the case. Phonological manipulation appears to incorporate the skills tapped by other phonological awareness tasks. To do a deletion task (e.g., to change *sneak* to *seek*), or a substitution task (go from *roof* to *room*), it appears that the student must be able to segment, isolate, and blend. Thus, to delete the /n/ from *sneak*, the student must separate the sounds (*segmentation)*, then *isolate* where the /n/ is located in the word, delete it (*manipulation*), and *blend* the remaining parts to arrive at *seek.* This suggests that manipulation tasks cast a wider net in terms of capturing more of the metalinguistic underpinnings associated with the construct of phonological awareness (Kilpatrick, 2012).

Given the current popularity of segmentation tasks, a manipulation task such as the PAST could be used to provide an additional or even more valid assessment phonological awareness construct as it relates to word level reading. For that very purpose, the PAST is here being offered to researchers and educational evaluators as a public domain assessment.

A Brief History of the PAST

As mentioned, the *PAST* began its life as Rosner and Simon's *AAT* (Rosner & Simon, 1971). Jerome Rosner worked closely with McInnis (1981, 1999) in training teachers in Rosner's *Auditory-Motor Program* (Rosner, 1974), which was a phonological training program that paralleled the *AAT* assessment. According to its copyright page, Rosner's *Auditory-Motor Program* became public domain in 1984, and by that point, McInnis had already expanded both the Rosner training program and his version of the *AAT*, which McInnis dubbed the *Language Processing Assessment,* (later changed to the *Phonological Processing Assessment,* or *PPA;* McInnis, 1999). Before researchers and educators settled on the use of the term "phonological awareness" (Scarborough & Brady, 2002), this skill was referred to by various terms, including *auditory analysis skill* (Rosner & Simon, 1971), *language processing* (McInnis, 1981), *linguistic awareness* (Blachman, 1984), *metalinguistic awareness* (Warren-Leubecker & Carter, 1988), as well as *phonological* and *phonemic awareness* (e.g., Lewkowicz, 1980). McInnis updated Rosner's training program as well as the *AAT* test based upon progress in research in phonological awareness. His updates included the recognition of differing levels of linguistic complexity within phonological awareness, (i.e., syllable, onset-rime, and phoneme levels, *cf.*

Treiman, 1991) as well as differing levels of difficulty of phonological awareness tasks based on the position of the target sound within a word (e.g., Stahl & Murray, 1994). In addition, he expanded the *AAT* by adding substitution items, which others have done as well (Lenchner, et al., 1990; Nikolopoulos, Goulandris, Hulme, & Snowling, 2006).

McInnis' PPA was used in schools for nearly three decades but was never the subject of any rigorous data collection effort. The *PAST* represents a revision of McInnis' *PPA*, making it, in a sense, a third generation of the Rosner and Simon AAT. Thus, the acronym "PAST" has double meaning, first as an acronym for *Phonological Awareness Screening Test*, but second as an acknowledgement that it has a long history, built upon the "past" work of its predecessors.

The Distinctive Aspects of the PAST

The *PAST* revised the *PPA* in three ways. First, it adds items involving manipulations that are not found on the *PPA*, such as deleting ending sounds and deleting and substituting penultimate sounds in single syllable words ending in a blend (e.g., going from *lift* to *list* by changing a /f/ to /s/). This was done because Ehri's theory of sight word recognition seems to imply that if students were phonemically aware of every sound in every position within spoken words, they would be at an advantage when they are learning to read (Ehri, 2005; Kilpatrick, 2013).

Second, the *PAST* provides feedback for every incorrect item. In at least one of their studies, Hulme and colleagues successfully used this practice with their phonological deletion task (Hulme, et al, 2002). Bryant (2002) criticized this practice as "unusual" with the claim that it turned the test into a "phonological training task." This is not likely. The corrective feedback given on the *PAST* acknowledges that phonological awareness typically takes 1-3 years to fully develop in normal readers, and therefore assumes that a student is not going to develop phonological awareness skills in the six to eight minutes it takes to administer the *PAST.* On the other hand, phonological awareness tasks often represent an unfamiliar set of expectations for students, and therefore a student may not have a clear understanding of exactly what the examiner wants him or her to do. This appears to be the case in some instances on the CTOPP Elision subtest. This subtest has a shift in task demands part way through the test without further instruction or feedback, and many students ceiling out at that point (Kilpatrick, 2012). The question becomes whether the student lacks the proper phonemic awareness to complete those items, or if instead he or she failed to adapt to the changing task expectations. With the PAST, it is presumed that providing feedback for every incorrect item might more adequately distinguish between students who lack phonological awareness skills from students who do not have a clear and sustained understanding of the task requirements. It is assumed that a student who starts improving his or her responses as a result of feedback likely possesses phonological awareness skills, but had previously not been clear on the task demands. By contrast, a student who fails to demonstrate phonological awareness skills despite feedback throughout this test is more likely to have a genuine lack of phonological awareness. It must be noted that this feature of the PAST, that is, providing feedback for every incorrect item, is not without precedent among phonological awareness tests. As mentioned above, Hulme and colleagues successfully used this approach on their experimenter-designed test. In addition, the *Segmenting Words* subtest from the *CTOPP* provides feedback on every incorrect item (Wagner, Torgesen, & Rashotte, 1999). In the highly influential study by Vellutino et al., (1996), the researchers used a procedure similar to this to insure they were assessing the students' phonological awareness skills and avoiding confusion with task demands. Bentin & Leshem (1993) gave kindergarteners 7 phonological awareness tests. Each test had five practice items, but more could be given if the experimenter felt that the student did not understand the task, addressing the issue of task demands in a different manner. No feedback was given for the actual items. However, given that the PAST has seven levels that are used in the scoring and five items per level, then Bentin & Leshem's procedure provides an equivalent amount of feedback, except that the PAST uses all those items in the score while Bentin & Leshem did not. Also, Bentin & Leshem's study met the rigorous standards necessary to be included in the National Reading Panel's (2002) meta-analysis. These illustrations should suggest that the practice of feedback for all items on the PAST, though not typical, is not without precedent. Finally it will be demonstrated below that the PAST displays correlations with tests of word identification and phonic decoding that are typically higher than commonly used phonological awareness tasks. This suggests that it is unlikely that the feedback feature of the *PAST* decreases its concurrent validity with reading. Rather, it may partially account for its superior correlations with word-level reading tasks.

The third distinctive feature of the PAST is its timing element and associated dual scoring system. When an item is given, the examiner mentally counts "one thousand one, one thousand two." If the student responds to that item before the examiner reaches the word "two" in the silently counted phrase, the student receives an automatic score for that item. If he or she responds after the examiner silently completes that phrase, the student receives a score as correct, but not automatic. All correct and automatic responses count toward the student's total correct score. However, only those items that were responded to within two seconds count toward the student's automatic score.

It would seem that such a mental count would introduce a great deal of error variance. However, there are two reasons this is not likely. First, school psychologists have been trained to properly provide a one second count per digit on the Wechsler Digit Span subtest. That test has demonstrated reliability and validity. Second, and more importantly, it is not common for an examiner to have to make a judgment as to whether to score the item as automatic or not. Based on our clinical experience with hundreds of students, an automatic response takes about one second, so a two second cutoff is rather generous. However, students who do not respond immediately typically take three to five seconds or longer to respond. In those cases, it is presumed that the student is strategizing, such as mentally spelling the word, manipulating the letter/sound as requested, and mentally "reading back" the result. Such a strategy represents a mental application of phonics, and does not necessarily reflect developed phoneme awareness.

The first author previously administered the McInnis PPA with older students referred for reading difficulties and found many of them could correctly respond to the items, but often took several seconds to complete a given item. Then, after field-testing the PPA with some younger students who were typical readers, he was struck by how these students tended to respond instantly, despite the fact that that these students had no previous exposure to this type of task. This prompted the inclusion of the timing element into the PPA/PAST in an attempt to distinguish students with phoneme awareness from those who could respond correctly on a phoneme awareness task by using some compensating mental/phonic strategy. Experience with first and second graders (from Study 1) made it clear that at this younger level, the automatic element had little impact on the interpretation of the test. At these grade levels, particularly at Grade 1, students either responded incorrectly or correctly and instantly. It is surmised that these students did not have the necessary spelling, phonic and/or metalinguistic sophistication to correctly use a strategy to accomplish such a task in a compensating manner. So, either they had the phonological awareness skills sufficient to respond instantly, or they did not. While most typical first and second grade readers do not appear to be able to use a slower, compensating strategy to arrive at the correct answer, older students seem to be able to do this.

It appears from both clinical experience and research data gathering experience with hundreds of students that older students with reading difficulties are often able to get items correct on a phonemic awareness test, using non-instantaneous responding. This creates the impression that phonemic awareness is not an issue for them because slow responses and instant responses count the same in raw scores, which are used when consulting a normative table. However, their responses are very slow and presumably strategic, and do not appear to be similar to the virtually effortless and instantaneous responses of younger, typical readers. This leads to an important problem. If poor readers are being hindered by inadequate phonemic awareness, it is possible in many cases that these students can perform adequately on conventionally scored phonemic awareness tests because those tests assign a correct answer to each item regardless of how quickly the student responds. In such cases it may be incorrectly inferred that the student does not have any phonological awareness difficulties. This is because the student's raw score translates in the norm tables to a score within a typical range of performance. In such cases a student's phonemic awareness difficulties will not be recognized or addressed.

Curriculum Based Measurements (CBM) of phonological segmentation (e.g., DIBELS, AIMSweb, easyCBM) use timing to some degree by assessing the number of correct responses within a specified time, usually one minute. This appears to us to be a valid idea. However, the CBM assessments rely on segmentation, which is a type of task that is less strongly correlated with reading, and which appears to lose its discriminant validity reading after first grade (Kilpatrick, 2012; Vloedgraven & Verhoeven, 2009; Wagner et al., 1993). Also, these CBM tasks are not used in their respective batteries past first grade, so they are not used at an age level in which the timing element would be most helpful. Despite these concerns, the timing element within the CBM based segmentation tasks appears to be a step in the right direction, but a timing element should be added to manipulation tasks, which are more strongly associated with reading development (Kilpatrick, 2012; Lenchner, et al., 1990; Swank & Catts 1994). In this article, we begin to explore the value of the timing element on the PAST.

Summary of Empirical Data on the PAST

While the LPA/PPA, which forms the immediate basis for the PAST, has been used since the late 1970s, The PAST itself, in its current form, has been used on a limited basis in schools since 2003, and we have collected data on this measure at various grade levels, from kindergarten to college, and with both typical readers as well as students with reading difficulties and disabilities. These data are presented below. The following data sets are presented separately in the traditional format of Study 1, Study 2, etc., though these primarily represent data gathering on comparing the performance on the PAST with word level reading or with other phonological awareness tasks. Because the PAST was used in each of the following studies, and will be presented in detail in the latter part of this article, it is not separately listed or described in any of the studies.

Study 1

Method

Participants. Participants included 67 unselected first grade students (30 female, 37 male) and 50 unselected second grade students (24 female, 26 male) from a lower middle class school district, in which 94% of students were white.

Materials. 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 read a graded word list. The *Word Attack* subtest involves sounding out pseudowords (e.g., *seeg, trast*). The participants' scores on these tests were the total number of items read correctly. In addition, students were administered three subtests from the *CTOPP; Elision, Blending Words,* and *Segmenting Words. Elision,* as mentioned above, is a modern derivative of the *AAT* and involves deleting a syllable, an onset, or a phoneme. *Blending Words* is a blending task starting with syllable blending and progressing to onset-rime blending and finally phoneme level blending. The student hears a word one part at a time (e.g., *can-dy*) and must indicate the word that the examiner is trying to say. *Segmenting Words* is a phoneme segmentation task, starting with two phoneme words and progressing to longer words (e.g., "What are the sounds in *bat*?" = /b/ /a/ /t/). *Memory for Digits* is a simple digits forward short-term memory task in which the student immediately repeats back the digits presented.

Procedure. The CTOPP tests were administered in normal CTOPP order (*Elision, Blending Words, Segmenting Words*) to preserve the order actually used by practitioners so as not to affect any normative comparisons (see more on order of administration below). These were followed by the WRMT-R tests, and finally the PAST. Some of the data from Study 1 were published as part of another project (citation, date [removed for blind review]) that did not make use of any of the PAST data presented below.

Results and Discussion

First Grade. Means and standard deviations for each subtest are included in Table 1 while the results from the correlational analysis are reported in Table 2. The PAST correlated with the

WRMT-R *Word Identification* r = .79 and *Word Attack* r =.81. These were somewhat higher than the correlations with *Elision* and *Blending Words,* and substantially higher than *Segmenting Words.* A hierarchical multiple regression analysis revealed that with these first graders, only the *Elision* subtest accounted for any unique variance beyond the PAST (Table 3). This is interesting because *Elision* and the *PAST* are both manipulation tasks. However, this result may be explained by some differences in task expectations. The *PAST* includes substitution items while *Elision* does not, and the *Elision* subtest has items that involve deleting phonemes from within two syllable words (e.g., saying *tiger* without /g/ = *tire*) while the *PAST* does not. The findings with *Word Attack* were nearly the same as with Word Identification, with only slight variations in magnitude.

Second Grade. For the second graders, the PAST correlated r = .76 for *Word Identification* and r = .83 for *Word Attack*. As in the first grade results, this was higher than *Elision* and *Blending Words* and substantially higher than *Segmenting Words*. Regression analyses indicate that only *Blending Words* contributed unique variance to *Word Identification* beyond the *PAST*. Why this finding differed from the first grade sample is not clear. However, it may be instructive to note that the correlation between *Blending Words* and the *PAST* was higher in the first grade sample (r = .65) than the second grade sample (r = .47). This suggests that between the first and second grade levels of these skills, there may be some degree of disengagement of either psycholinguistic or task-level requirements that result in these two tasks each accounting for unique variance.

Wagner & Torgesen (1987) referred to Lundberg, et al.'s (1980) correlation of r = .75 between reading and phoneme reversal as "astonishingly high." From this statement, it can be inferred that the *PAST* yields correlations with word-level reading achievement with this particular sample of first and second graders that could be described as "astonishingly high." However, correlations this high are not unprecedented for phonological manipulation tasks. The Lundberg, et al. (1980) study just mentioned illustrates this. In addition, Lenchner, et al., (1990) had very similar strengths of correlation with reading on their manipulation task (.78 & .79). Also, Rosner & Simon (1971) found a correlation of .84 between the *AAT* and their reading measure at the third grade level, and as previously stated, the *PAST* originated with the *AAT*.

If the *PAST* and the *Elision* subtest are both manipulation tasks derived from the AAT, why the difference in correlation? An examination of the scatter plots (Figure 1) and of the task items may answer this question. The *Elision* subtest has a change in task requirements at Item 9, but all feedback stops after Item 5. Items 9-11 require deleting a phoneme from the middle of two syllable words (e.g., "Say *tiger* without saying /g/"). Clinical observations during testing indicated that many students respond *tie* to Item 9, and no corrective feedback was allowed. Given that the ceiling on the subtest is three items incorrect in a row, it was common for students to get 9-11 incorrect, which ends the test. The scatter plots for the Elision subtest for both grade levels are consistent with this pattern. None of the other three subtests had a similar pattern, namely numerous relatively stronger readers performing at or below median on their respective phonological awareness test. This is consistent with the possibility that it was a lack of clarity with the shift in task demands that several of the stronger readers did not advance beyond Item 9, and suggests this may not be an accurate reflection of their actual phonological awareness skills. Thus, this administration feature of the *Elision* may partially account for the difference between these two manipulation tasks, and presumably, it may be speculated that if there were feedback given for every item (or at least when the new task demands were introduced), the strength of correlation between *Elision* and reading would have been greater and there would not likely have been such a “spike” in the upper left quadrant of the scatter plot.

One further consideration is order of administration. The CTOPP subtests were administered in that battery’s standard order. The PAST was administered afterward so that exposure to a task similar to the Elision subtest would not influence the students’ performance on that latter task, given that the CTOPP normative sample would not have had a phonological manipulation task beforehand. However, the fact that the PAST was administered after other phonological awareness tasks leaves open the possibility that the very high correlations were due to the fact that the students were already “warmed up” for this type of task, eliciting better phonological awareness performance than would otherwise be the case. While this possibility cannot be ruled out, it is not altogether clear precisely how the PAST being administered subsequent to the Elision subtest resulted in stronger correlations between the PAST and the reading measures. Presumably they would be more attuned to this type of task and thus the last test being a more valid assessment of a student’s phonological awareness skills.

An examination of the scatter plots in Figure 1 demonstrates an interesting difference between the two manipulation tasks (*Elision* and *PAST*) on the one hand and the *Blending Words* and *Segmenting Words* subtests on the other. The latter two appear to have a substantial number of first grade participants who did well on the blending and segmenting tasks yet poorly on the reading measure. By contrast, none of the weaker readers did well on either of the manipulation subtests. There are at least two possible ways of interpreting this. The first would be to simply suggest that these manipulation tasks are more difficult than the blending and segmenting tasks. Various researchers have demonstrated that manipulation tasks are too difficult for most kindergarteners while segmentation and blending are not (e.g., Ball & Blachman, 1991;

Stanovich, et al., 1984; Wagoner, et al., 1993; Yopp, 1988; more on this below). However, these participants were first-graders and it is clear that many of them were able to do the manipulation tasks. The stronger readers displayed no apparent difficulties while it was only the weaker readers who did poorly on the manipulation tasks. A second possibility is that the segmenting and blending tasks are not as sensitive to reading difficulties as are the manipulation tasks. Previous research provides data to support this (Kilpatrick, 2012; Swank & Catts, 1994).

One of the important findings from Study 1 is that the PAST displays strong concurrent validity with both the reading measures and the other phonological awareness subtests from the CTOPP (except Segmenting Words in second grade). Indeed, the PAST correlated more strongly with the reading measures than any of the three CTOPP phonological awareness subtests at both grade levels, providing substantial support for its validity. However, as mentioned, the order of administration may have influenced these outcomes.

To get a sense of whether the order of items within of the PAST genuinely represent an increase in difficulty as the test progresses, mean correct scores for each level of the PAST are provided in Table 4. It can be seen from this that there is a definite trend. As the test progresses, the items generally become more difficult. Exceptions to this are that the onset-rime levels had higher scores than one of the syllable levels, Level E. Three of the five items in Level E involve removal of a stressed syllable in a three syllable word (e.g., “Say *barbecue* but don’t say *bar.*”). This task appears to be more challenging than simply removing or exchanging the first sound in a single syllable word (Level F: “Say *sand* without the /s/”; Level G: “Say *make* and change the /m/ to /l/”). Within the phoneme levels, some of the levels were somewhat easier than the level before. However, the magnitude of these differences on the PAST were not sufficient to warrant a change in label/order (i.e., J, K, L) compared to the PPA, the PAST’s immediate predecessor.

The PPA used three decades of clinical data to establish its order. Here we are reporting on two grade levels from one school. With that said, it is clear that as one progresses from D through M, there is a marked tendency toward more difficult items (which is even more pronounced when the automatic scoring was used). Further analysis of other data sets may provide additional clarification and possible adjustments.

It must be noted that on the PAST, a ceiling is reached when a student gets a zero or one out of five at two consecutive levels. However, it is likely that this procedure affected the finding in Table 4 because means are used and it can be seen that for those who continued on further in the test, there was an increasing level of difficulty. Thus, it may seem fair to assume that those who hit ceiling earlier in the test would have found the higher levels more difficult. Second, while these results do not represent anything like national norms, it may provide the reader with a very broad general framework for what to expect in a general education, lower middle class first and second grade sample (data gathered between December to March) in terms of skills at the various levels on the PAST. Table 1 already provided the overall means and standard deviations of the test.

Study 2

Method

Participants. Participants included 60 unselected second grade students (35 female, 25 male) and 67 unselected fifth grade students (32 female, 35 male) from the same school district as Study 1.

Materials. All participants received the *Word Identification* and *Word Attack* from the *Woodcock Reading Mastery Test, Revised* (*WRMT-R*; Woodcock, 1999). They were also administered the *Exceptions Word Test* (*EWT*; Adams & Huggins, 1985). The EWT is a graded word list in which all words include one or more exceptions to standard grapho-phonemic pronunciation rules (e.g., *one, colonel*). The *Spelling* and *Reading Comprehension* subtests from the *Wechsler Individual Achievement Test* (WIAT) were administered. *Spelling* requires students to spell progressively more difficult words, while in *Reading Comprehension,* students answer questions about passages they read (Wechsler, 1991). The participants' scores on all these tests were the total number of items read, spelled, or answered correctly. In addition, the *Blending Words* subtest from the CTOPP was also administered. A subset of 32 of the second graders were administered the *Phoneme Segmentation* subtest from the *Developmental Indicators of Basic Literacy Skills* (DIBELS). The data in Study 2 was collected in two cohorts, the first in the spring of 2008 and the second in the spring of 2010. The DIBELS Phoneme Segmentation subtest was included in the 2008 battery. However, it had such a poor showing in terms of correlations with other measures, it was not included in the 2010 data collection. Thus, the subset of 32 participants represents the second graders from the 2008 cohort.

Results and Discussion

The means and standard deviations for each subtest at both grade levels are reported in Table 5. The results from a correlational analysis of the subtests are presented in Tables 6 and 7. With the second graders, the PAST correlated with the WRMT-R *Word Identification* $r = .57$; and *Word Attack* $r = .57$, *Reading Comprehension* $r = .62$, *EWT* $r = .44$, and .48 for WIAT *Spelling*. For the fifth graders, the PAST had a significant correlation with *Word Identification* ($r = .29$), *Word Attack* ($r = .30$), the EWT ($r = .38$) and WIAT *Spelling* ($r = .37$), but did not significantly correlate with *Reading Comprehension* ($r = .17$). For the subset of 32 second graders administered the *Phoneme Segmentation* from the DIBELS, there were no significant correlations between the DIBELS Phoneme Segmentation subtest and the other measures, except for a significant negative correlation with the PAST.[1] This data is presented in Table 7. It must be made clear that the DIBELS Phoneme Segmentation subtest is used only with first graders in the DIBELS battery, so our findings in no way reflect on its usage with first graders. However, when considering the DIBELS Phoneme Segmentation subtest as a representative of a phonological segmentation task, our results are consistent with the findings of others (Kilpatrick, 2012; Vloedgraven & Verhoeven, 2009; Wagner et al., 1993). It appears that phonological segmentation loses its discriminant validity with reading after first grade, while other phonological awareness tasks, such as manipulation, continue to provide an index of the phonological correlates of reading beyond first grade (Kilpatrick, 2012; Vloedgraven & Verhoeven, 2009; Wagner et al., 1993).

Study 2 confirms the results from Study 1 in that the PAST displays a significant correlation with word-level reading tasks. It extended the Study 1 results in a few ways, such as demonstrating significant correlations with exception word reading and spelling at both grade levels, and reading comprehension at second grade. There was, however, a notable disparity in strength of correlation between the PAST and the reading measures with the second grade sample in this study compared to the second graders in Study 1. This disparity is all the more interesting because the Study 2 data were gathered from the same school building as the Study 1

[1]Even though the Phoneme Segmentation subtest is timed, higher scores are still better (unlike timed tests of rapid naming, where lower scores are better). Thus, the expectation is for positive correlations with measures of reading and other phonological awareness tests. The significant negative correlation with the PAST can be accounted for based upon an observed tendency among the students doing the test. Students who did better on the PAST seemed to display a tendency, based on a desire to respond quickly, to default to an onset-rime manner of segmentation, which lowered their scores. Because no feedback is allowed once the timed test begins, the stronger students did not seem to be aware they were not doing it correctly.

data, except gathered in 2008 and 2010, while the data in Study 1 were collected in 2004. While the reasons for this disparity are by no means certain, one consideration relates to the characteristics of the samples of second graders. A direct comparison of the overlapping tests administered is presented in Table 8. The first difference between the groups is that Study 1 represents archival data from a “universal screening” of every student in their respective first and second grade classrooms conducted for the benefit of the reading teachers and classroom teachers for instructional purposes. The participants in Study 2 were involved as research participants and the only students who participated were those for whom we received student assent and parental consent. Thus, a selection bias could have the make up of the groups. Second, it can be observed in Table 8, the mean performance on all of the subtests except Word Attack were higher with the second graders in Study 2. But what may be more important is that the standard deviation in the Study 2 second graders’ Word Identification subtest was half as large as their Study 1 counterparts. Also, their standard deviation on the PAST was somewhat lower. These differences support the idea that less variability in the Study 2 sample resulted in the lower correlations. Finally, the Study 1 data were gathered between December and March, while the data for Study 2 were gathered from April to June. While that fact does not account for the lack of “growth” on the Word Attack subtest, it may partially account for the differences on the PAST and the Word Identification subtest. As we will argue below, typical second grader readers are approaching a level of phonological awareness development that levels off and thus would have a decreasing correlation with word level reading skills. The average of a few months difference between these groups may have made a difference here. Thus, the few months difference in age, the potential selection bias, and the greater variation in performance within the samples may account for the differences between these two groups of second graders.

Fifth Grade Though significant, the correlations between the PAST and the reading measures were much weaker than the correlations in Study 1 as well as with the second graders in this study. This is likely due to the fact that by fifth grade, typical readers have developed the necessary phonological awareness skills to read, and this near ceiling performance reduces variability, and thus the correlations. There were very few low readers in this sample (only 7% of students were below the 25th percentile on the WRMT-R Word Identification). Thus, this group had typical reading and phonological awareness, and by fifth grade, typical phonological awareness is near ceiling on the PAST (see Study 3).

It was previously mentioned that the PAST has a dual scoring system, with one score representing the number correct while the second score represents the number of items correct that were responded to in two seconds or less. An examination of Table 6 indicates that in the second grade sample, the correlation between WID and the PAST-A (i.e., the automatic score) is lower than the correlation with the standard, "items correct" score on the PAST (.52 vs. .57). This is also true for the correlations with the Exceptions Word Test (.41 vs. .44), as well as an even more substantial difference for the Word Attack subtest (.36 vs. .57). However, with the fifth grade students, this trend is reversed. For all three measures of reading, Word Identification (.37 vs. .29), Word Attack (.36 vs. .30), and the Exceptions Word Test (.47 vs. 38), the automatic score correlations were higher than the untimed score. What may be of interest is that in second grade, the PAST & PAST-A correlated least with EWT compared to the WID and WA (except for the PAST-A with WA), but in fifth grade, both the PAST and PAST-A correlated the most strongly with Exception Words Test. Of those three subtests, the Exception Words Test presumably assesses the most developed phase of word-level reading (Ehri 2005). The implication is that those with the strongest phonemic awareness scores, as tested on the PAST, are apparently the ones with the strongest sight vocabularies.

In terms of word level reading, these three subtests like along a natural continuum. On one end is Word Attack, which is based on reading nonsense words and can only be successfully completed using phonic decoding skills. On the other end is the Exception Words Test, which contains words that do not follow a strictly phonic pattern and readers are most successful if they are already familiar with the specific words because by itself, a phonic decoding strategy will not yield the correct response. In between these two is Word Identification in which correct responses can result from either prior familiarity with the words (akin to the EWT) or by sounding out the phonetically regular words on that test (akin to Word Attack). The correlational strength of the PAST results lie along this continuum. The implication is that by fifth grade, the phoneme awareness assessed on the PAST is more strongly associated with establishing words as familiar for later recognition (i.e., a student's pool of sight words) as reflected on performance on the Exception Words Test, and less associated with phonic decoding. This pattern seems even clearer when we compare the fifth grade results with the second grade results. Sight vocabulary and orthographic memory is built upon a foundation of phonemic awareness and sound symbol skills (Ehri, 2005; Kilpatrick, 2013; Share, 1995), and the trend in our findings is consistent with that development. It would not be appropriate, however, to make too much of this pattern given the disproportionately small percentage of weak readers in the data set. But it would appear that considering the automaticity of phonological awareness skills is an area that may be worth further study.

Because of the characteristics of this particular sample of fifth graders, it would not be expected that there would be many students who were weak readers with poor phonological awareness who compensate their way to an acceptable score on a phonological awareness test via strategizing. However, even with this said, the correlations tend in the direction previously mentioned, that older students may have the ability to compensate, while the younger ones are less likely to compensate. Because any general population containing a representative sample of typical, older readers is not likely to contain many compensators, as we have conceptualized them, a group statistic may mask the existence of compensators. The fact that there was a shift in correlational strength in the automatic scoring of the PAST between the second and fifth grade samples may be more significant than is obviously apparent. To examine more individualized performances to determine if this compensating trend exists, scatter plots were generated with both second and fifth graders.

Figure 2 visually displays the tendency we are describing. When examining the scatter plots of the second graders, there seems to be very little overall difference in the standard vs. automatic scoring of the PAST, other than overall lower scores as a group. However, there may be a little tendency for the upper left of the scatter plot to become more populated as we shift from items correct to items automatic, suggesting that some of the stronger second grade readers did not have automatic phonological awareness skills, or were simply more hesitant to respond because they wanted to be correct (at no point do the instructions ask the students to respond quickly). However, in the fifth grade sample, the apparent compensating tendency seemed to emerge somewhat. There were students in the lower right of the scatter plot who “disappear” when the automatic scoring is used. This is consistent with the suggestion that some weaker readers may do well on a phonological awareness test via compensation, yet when automaticity is taken into account, these weaker readers no longer display adequate phonological awareness. This is a phenomenon that would not be captured by a group statistic, which is the most common level of analysis in the phonological awareness literature. Because of the underrepresentation of weak readers in this sample, the magnitude of this tendency is likely to be limited. However, it will be seen below that with a sample of students referred for reading difficulties, the pattern is more apparent (see Study 6 and 7 below).

Study 3

Method

Participants. Participants included 30 college students (20 female, 10 male) from a state university college in Upstate New York who were recruited from an introductory psychology class. They ranged in age from 18 to 22 years old (mean age was 19.3 years). Twenty-six of the students were of white, European decent, three were African American, and one was Hispanic. Students provided informed consent and received partial course credit for participating. No pre-selection criteria were established ahead of time related to a student's reading level or a potential reading disability.

Materials. All participants received the *Word Identification* and *Word Attack* subtests from the *Woodcock Reading Mastery Test, Revised* (*WRMT-R*; Woodcock, 1987, 1999).

Results and Discussion

Means and standard deviations are presented in Table 9. Table 10 provides the results from the correlational analysis. Among the adults in Study 3, the *PAST* correlated with the *WRMT-R Word Identification* r = .54 and with Word Attack r = .53. This is higher than most segmentation tests correlate with reading among young children. The significance of this will be addressed in the general discussion. What is interesting is that the correlation between the PAST and the WRMT-R Word Identification was stronger among the college students than among the fifth graders in Study 2. Table 11 directly compares the performance of the fifth graders and the college students. What may be most interesting is that there is virtually no difference between college students and fifth graders on the PAST's untimed score. However, when the automatic scoring was used, the college students displayed higher scores than the fifth graders, suggesting a greater level of automatization of their phonological awareness skills. This suggests that while phonological awareness seems to have leveled out to an adult level by fifth grade in terms of untimed responding, there continues to be a fine tuning of this skill in terms of automatizing that continues beyond fifth grade. The degree to which this may affect later reading development, if it does at all, is a question that cannot be answered by our current data. Others have found that adult readers with reading difficulties continued to have phonemic awareness difficulties (Bruck, 1992; Greenberg, et al., 1997). Thus, based upon these results, the PAST appears to have the potential to serve as a tool that can distinguish among stronger and weaker adult readers in terms of their level of phonemic awareness skills.

Study 4

Method

Participants. Participants included 58 kindergarten students (31female, 27 male) from a lower middle class school district different from the school district used in Study 1 and 2.

Materials & Procedures. All participants received several subtests from the AIMSweb assessment system, including *Letter Name Fluency, Letter Sound Fluency, Phoneme Segmentation Fluency,* and *Nonword Fluency.* All of these tests were administered twice, once in March of kindergarten, and a second in June. All students were screened to determine eligibility for a volunteer afterschool program and to pilot the AIMSweb assessment system in that school. The assessment team members chose to include the PAST in their universal screening.

Results and Discussion

The means and standard deviations on all subtests are listed in Table 12. The intercorrelations from each of the measures are reported in Table 13. Study 4 allows for an examination of the PAST's reliability due to multiple administrations two and a half months apart. While good reliability can be inferred based upon the strong validity data reported in Studies 1, 2, and 3 above, Study 4 allows for a more direct measure. The multiple administrations, however, involved alternative forms of the PAST. The only difference in the forms was the actual words used as stimulus prompts. So, there is a natural confound of test-retest reliability and alternative form reliability. Despite the measurement error inherent in confounding time of administration and alternate forms, the correlation between the first and second administrations of the PAST was $r = .82$. This was higher than the AIMSweb Letter Sound Fluency ($r = .73$) and Phoneme Segmentation Fluency ($r = .72$) subtests, and equal to Letter Name Fluency ($r = .82$). Only Nonword Fluency was higher ($r = .89$). Thus, the PAST displayed reliability equivalent to or higher than most of the AIMSweb subtests.

From Table 13 it is clear that there are very strong intercorrelations among all of the measures of early literacy. For our purposes, it is interesting to note that the only index of reading, Nonword Fluency, was more highly correlated with the PAST at first administration ($r = .65$) than the AIMSweb's own measure of phonological awareness, Phoneme Segmentation Fluency ($r = .43$). A hierarchical linear regression **[ADD TABLE]** indicated that with *Nonword Fluency* as the dependant variable, *Phoneme Segmentation Fluency* accounted for no variance beyond the variance captured by the PAST test. These results are similar to previous findings comparing phonological manipulation and phonological segmentation with reading (Kilpatrick, 2012; Swank & Catts, 1994). It is also interesting to note that Phoneme Segmentation Frequency did not account for any unique variance in Nonword Fluency performance in this group beyond what was found in the Letter Name Fluency subtest. By contrast, the PAST and Letter Name Fluency each contributed independent and highly significant variance in Nonword Reading Fluency.

It was mentioned earlier that previous researchers demonstrated that manipulation tasks were too difficult at kindergarten (Rosner & Simon, 1971; Stanovich et al., 1984; Wagner, et al., 1993; Yopp, 1988). Indeed, Wagner and colleagues (1993) found that in their kindergarten sample, segmentation had a stronger correlation with reading ($r = .38$) than manipulation ($r = .18$) (this was substantially reversed in their second grade sample with manipulation correlating with reading $r = .51$ and segmentation correlating $r = .27$). The difference in these previous findings with our current results is that those previous studies had kindergarteners manipulate at the onset-rime and/or phoneme levels. By contrast, the PAST uses syllable deletion at the start of the test (e.g., "Say *birthday* without saying *birth*"; "Say *enter* without saying *ter*"), which better distinguishes students at the lower end developmentally, because using onset-rime or phoneme level deletion in kindergarten results in floor effects (which Wagner et al., 1993 acknowledged). It is interesting that the CTOPP Elision has addressed this problem by adding manipulation items at the syllable level at the beginning of that subtest. So, it appears that the older view that manipulation tasks are too difficult for kindergarteners has confounded task and level. It may not be that the manipulation task, per se, is too difficult. Rather it seems that the onset-rime and phoneme levels are too difficult for many or even most kindergarteners. Thus, the PAST appears to be appropriate for kindergarteners because of its syllable level items. Indeed, it had a much stronger showing than the phonological awareness subtest on the current AIMSweb battery.

Study 5

Method

Participants. Participants included 17 fifth graders (10 female, 7 male) from Study 2 who coincidentally were first graders whose data are included in Study 1. This allows for an examination of the predictive validity of the measures in Study 1, albeit with a small sample. All students who were in both samples were invited to participate. Three additional students from the original first grade sample who were available in fifth grade at the time of Study 2 either did not return a parental permission form or chose not to sign a student assent to participate.

Materials. See Study 1 and Study 2.

Results and Discussion

Table 14 displays the intercorrelations between the tests administered between first and fifth grade. It can be seen that the PAST administered in first grade correlated with fifth grade word level reading (r = .60 with the WRMT-R Word Identification and r = .53 for Word Attack). These correlations were significant. Due to the limited sample size, the CTOPP Elision subtest from first grade was only marginally significant with fifth grade Word Identification (r = .47; p = .057), while neither CTOPP Blending Words (r = .29) nor Segmenting Words (r = .23) were significantly correlated with fifth grade Word Identification. While Blending Words administered in first grade failed to significantly correlate with Word Attack in fifth grade, Segmenting Words was marginally significant (r = .46; p = .064).

Given the sample size in Study 5 and the ad hoc nature of the data (i.e., an unintended and coincidental inclusion of similar students in different studies, five years apart), we must view these data with caution. However, the tendency in the data is clearly consistent with the other data collection efforts using the PAST. Study 5 appears to suggest that the PAST displays predictive validity that is consistent with the concurrent and construct validity presented in the other studies listed above in that it correlates moderately to strongly with word level reading skills and with other measures of phonological awareness.

Studies 6, 7 & 8

General Discussion

REFERENCES

Adams, M. J. & Huggins, A. W. F. (1985). The growth of children's sight vocabulary: A quick test with educational and theoretical implications. *Reading Research Quarterly, 20(3)*, 262-281.

Anthony, J. L., Lonigan, C. .J., Driscoll, K., Phillips, B. M., and 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.

Ball, E. W. & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26(1)*, 49-66.

Bentin, S., & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia, 43*, 125-148.

Blachman, B.A. (1984). Relationship of rapid naming ability and language analysis skills to kindergarten and first-grade reading achievement. *Journal of Educational Psychology, 76(4),* 610-622.

Bruce, D. J. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology, 34*, 158-170.

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

Bryant, P. (2002). It doesn't matter whether onset and rime predicts reading better than phoneme awareness does or vice versa. *Journal of Experimental Child Psychology, 82,* 41-46.

Cardoso-Martins, C., Mamede Resende, S., & Assunção Rodrigues, L. (2002). Letter name knowledge and the ability to learn to read by processing letter-phoneme relations in words: Evidence from Brazilian Portuguese-speaking children. *Reading and Writing: An Interdisciplinary Journal, 15(3-4),* 409-432.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research. 36,* 948-958.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3(4),* 331-361.

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 implementation. *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.

Dixon, M., Stuart, M. & Masterson, J. (2002). The relationship between phonological awareness and the development of orthographic representations. *Reading and Writing: An Interdisciplinary Journal, 15(3-4),* 295-316.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9(2)*, 167-188.

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(2),* 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.

Hulme, C., Hatcher, P. J. Nation, K., Brown, A., Adams, J. & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82,* 2-28.

Kilpatrick, D. A. (2012). Phonological segmentation assessment is not enough: A comparison of three phonological awareness tests with first and second graders. *Canadian Journal of School Psychology, 27(2)*, 150-165.

Kilpatrick, D. A. (2013). Tailoring interventions in reading based on emerging research on the development of word recognition skills. In Mascolo, J. T., Flanagan, D. P., & Alfonso, V. C. (Eds.) *Essentials of planning, selecting and tailoring intervention: Addressing the needs of the unique learner.* New York: Wiley.

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(4)*, 240-247.

Lewkowicz, N. K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72(5),* 686-700.

Lundberg, I., Olofsson, Å., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology, 21(3)*, 159-173.

McInnis, P. J. (1981). *Decoding keys for reading success.* New York NY: Walker Educational.

McInnis, P. J. (1999). *A guide to readiness and reading: Phonemic awareness and blending* (3rd ed.). Penn Yan, NY: ARL.

National Reading Panel. (2000). *Teaching children to read. The report of the National Reading Panel: Reports of the subgroups.* Washington, D.C.: US Government Printing Office.

Nikolopoulos, D., Goulandris, N., Hulme, C. & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94,* 1-17.

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.

Rosner, J. & Simon, D. P. (1971). The Auditory Analysis Test: An initial report. *Journal of Learning Disabilities, 4(7)*, 384-392.

Scarborough, H. S. & Brady, S. A. (2002). Toward a common terminology for talking about speech and reading: A glossary of the "phon" words and some related terms. *Journal of Literacy Research, 34(3),* 299-336.

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(3),* 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(2),* 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(2),* 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.

Treiman, R. (1991). The role of intrasyllabic units in learning to read. In Rieben, L. & Perfetti, C. A. (Eds.) *Learning to read: Basic research and its implications* (pp. 149-160). Hillsdale, NJ: Erlbaum.

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.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denkla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88,* 601-638.

Vloedgraven, J. & 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. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101,* 192-212.

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.

Warren-Leubecker, A. & Carter, B. W. (1988). Reading and growth in metalinguistic awareness: Relations to socioeconomic status and reading readiness skills. *Child Development, 59(3),* 728-742.

Wechsler, D. (1991). *The Wechsler Individual Achievement Test.* San Antonio, TX: Psych Corp.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23(3),* 159-177.

Hulme, C., Caravolas, M., Málková, G., & Brigstocke, S. (2005).

Table 1

Means and Standard Deviations for Study 1

Grades 1 & 2

	Grade 1 (*n* = 67) Mean (SD)	Grade 2 (*n* = 50) Mean (SD)
Age	6 yrs. 6 mos. (3.7 mos.)	7 yrs. 10 mos. (3.7 mos.)
WRMT-R Word Identification	28.18 (15.84)	50.88 (16.37)
WRMT-R Word Attack	11.79 (8.06)	22.56 (10.55)
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)
PAST	9.67 (7.37)	18.30 (7.27)

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing; PAST = Phonological Awareness Screening Test.*

Table 2

Subtest Intercorrelations for Study 1

Grade 1 (n = 67) and Grade 2 (n = 50)

	WID	WA	EL	SEG	BW	*PAST*
WID		.88***	.56**	.31*	.64**	.76**
WA	.81***		.67***	.33*	.51***	.83***
EL	.60***	.61***		.20	.29*	.63**
SEG	.47***	.44***	.26*		.35*	.26
BW	.65***	.59***	.49***	.55***		.47**
PAST	.79***	.81***	.58***	.36**	.65***	

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*; *PAST* = Phonological Awareness Screening Test.

*$p < .05$.

**$p < .01$

***$p < .001$

Table 3

Regression Analyses

Grade 1 (*n* = 67) and Grade 2 (*n* = 50)

Dependent Variable	Model	Independent Variables	Grade 1 β	Grade 1 R^2 change	Grade 1 *p* value	Grade 2 β	Grade 2 R^2 change	Grade 2 *p* value
WIAT-R *Word Identification*								
	1	CTOPP *Segmenting Words*	.47	.222	<.001	.31	.097	.029
	2	CTOPP *Segmenting Words*	.17	.019	.145 (ns)	.11	.011	.369 (ns)
		CTOPP *Blending Words*	.56	.215	<.001	.57	.282	<.001
	3	CTOPP *Segmenting Words*	.17	.020	.101 (ns)	.06	.003	.58 (ns)
		CTOPP *Blending Words*	.37	.078	.002	.48	.189	<.001
		CTOPP *Elision*	.38	.108	<.001	.42	.159	<.001
	4	CTOPP *Segmenting Words*	.16	.018	.057	.04	.001	.676 (ns)
		CTOPP *Blending Words*	.11	.005	.305 (ns)	.30	.067	.003
		CTOPP *Elision*	.18	.021	.041	.16	.008	.276 (ns)
		PAST	.56	.151	<.001	.57	.168	<.001

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing; PAST = Phonological Awareness Screening Test.*

Table 4

Mean Performances on Each Level of the PAST from Study 1

	Level	Grade 1		Grade 2	
		Scoring Approach		Scoring Approach	
		Correct	Automatic	Correct	Automatic
Syllable	D	4.3	3.9	4.9	4.7
Levels	E	3.6	3.2	4.6	4.1
Onset-Rime	F	4.9	4.8	5.0	4.8
Levels	G	4.5	4.2	4.9	4.4
Phoneme	H	2.6	2.0	3.7	2.8
Levels	I	2.9	1.5	4.2	2.4
	J	1.6	1.0	3.8	2.1
	K	1.7	0.7	2.7	1.0
	L	2.0	0.9	2.9	1.0
	M	1.3	0.4	2.4	0.6

Notes:
1) All raw scores reported above are out of a possible 5 points.
2) Because of the minimal variability in performance at the syllable and onset-rime levels (Levels D-G) and because only phoneme-level processing interfaces with reading from first grade and beyond (Ehri, 2005; Kilpatrick, 2013), only the phoneme levels were used in the calculations in the present paper (i.e., Levels H to M; highest possible total raw score = 30).

Table 5

Means and Standard Deviations for Study 2

Grades 2 & 5

	Grade 2 (n = 60)		Grade 5 (n = 67)	
	Mean (SD)		Mean (SD)	
WRMT-R Word Identification	53.53	(7.73)	73.06	(7.17)
WRMT-R Word Identification (SS)	**???**		101.68	(6.82)
WRMT-R Word Attack	21.69	(7.15)	29.91	(5.36)
WRMT-R Word Attack (SS)	**???**		102.98	(8.47)
Exception Words Test	17.20	(8.2)	37.98	(7.04)
WIAT Reading Comprehension	16.35	(3.85)	23.82	(5.95)
PAST	19.83	(5.43)	24.72	(5.04)
PAST automatic score	11.35	(5.78)	16.66	(5.48)
CTOPP Blending Words	15.11	(2.97)	16.03	(3.29)
WIAT Spelling	11.36	(2.29)	17.30	(3.50)

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing; PAST = Phonological Awareness Screening Test.*

FILL EMPTY CELLS ABOVE

Table 6

Subtest Intercorrelations for Study 2

Grade 2 (*n* = 60) and Grade 5 (*n* = 67)

	WID	WA	EWT	RC	PAST	PAST-A	BW	SPELL
WID		.61**	.86**	.41**	.29*	.37**	.10	.71**
WA	.74**		.61**	.28*	.30*	.36**	.25*	.68**
EWT	.81**	.58**		.44**	.38**	.47**	.09	.71**
RC	.79**	.58**	.73**		.18	.25	.29*	.28*
PAST	.57**	.57**	.44**	.62**		.78**	.06	.37**
PAST-A	.52**	.36**	.41**	.52**	.76**		.23	.54**
BW	.22	.22	.02	.22	.31*	.31*		.08
SPELL	.73**	.66**	.73**	.46**	.48**	.33*	.08	

Note: Grade 2 is below the diagonal; Grade 5 is above.

WID = WRMT-R *Word Identification*; WA = WRMT-R *Word Attack*; EWT = *Exception Words Test*; RC = WIAT *Reading Comprehension* subtest; PAST = *Phonological Awareness Screening Test*; BW = CTOPP *Blending Words*; SPELL = WIAT *Spelling* subtest.

*p < .05.

**p < .01

***p < .001

Table 7

Subtest Intercorrelations for the Second Graders who were

Administered the DIBELS Phoneme Segmentation Subtest

Grade 2 (n = 32)

	WID	WA	EWT	RC	PAST	BW
WA	.81**					
EWT	.74**	.55**				
RC	.78**	.41	.71**			
PAST	.43*	.43*	.51**	.34		
PS	-.18	-.09	-.26	-.11	-.36*	
BW	.51**	.42*	.16	.36*	.35*	
SPELL	.76**	.75**	.71**	.58**	.43*	.38*

WHY ISN'T 'PS' ACROSS THE TOP??

Note: WID = WRMT-R *Word Identification*; WA = WRMT-R *Word Attack*; EWT = *Exception Words Test*; RC = WIAT *Reading Comprehension* subtest; PAST = *Phonological Awareness Screening Test*; PS = DIBELS *Phoneme Segmentation* subtest; BW = CTOPP *Blending Words*; SPELL = WIAT *Spelling* subtest.

*p < .05.

**p < .01

Table 8

A Comparison of Performance Between

Second Graders in Study 1 and Study 2

	Study 1 (*n* = 50) Mean (SD)	Study 2 (*n* = 60) Mean (SD)
WRMT-R Word Identification	50.88 (16.37)	53.53 (7.73)
WRMT-R Word Attack	22.56 (10.55)	21.69 (7.15)
CTOPP Blending Words	14.24 (2.93)	15.11 (2.97)
PAST	18.30 (7.27)	19.83 (5.43)

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing; PAST = Phonological Awareness Screening Test.*

Table 9

Means and Standard Deviations for Study 3

(n = 30)

	Mean	SD
WRMT-R Word Identification	90.47	4.77
WRMT-R Word Attack	33.10	4.67
PAST	26.23	2.01
PAST Automatic scoring	21.00	4.77

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised; PAST = Phonological Awareness Screening Test.*

Table 11

A Comparison of Performance Between

Fifth Graders (Study 2) and College Students (Study 3)

	Grade 5 (n = 67) Mean (SD)		College Students (n = 30) Mean (SD)		Difference Raw (% increase)	
WRMT-R Word Identification	73.06	(7.17)	90.47	(4.77)	17.41	(23.8%)
WRMT-R Word Attack	29.91	(5.36)	33.10	(4.67)	3.19	(10.7%)
PAST	24.72	(5.04)	26.23	(2.01)	1.51	(6.1%)
PAST automatic score	16.66	(5.48)	21.00	(4.77)	4.34	(26.1%)

Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing; PAST = Phonological Awareness Screening Test.*

Table 12

Means and Standard Deviations for Study 4

(n = 58)

	Time 1 Mean (SD)	Time 2 Mean (SD)
Letter Name Fluency	40.57 (16.34)	41.79 (15.15)
Letter Sound Fluency	20.05 (10.77)	22.77 (9.12)
Phoneme Segmentation Fluency	27.18 (16.45)	34.59 (17.63)
Nonword Fluency	26.66 (25.17)	32.86 (29.18)
PAST	16.95 (9.99)	19.20 (10.46)

Table 13

Subtest Intercorrelations for Study 4

(n = 58)

	LNF1	LNF2	LSF1	LSF2	PSF1	PSF2	NWF1	NWF2	PAST1
LNF2	.82***								
LSF1	.65***	.66***							
LSF2	.69***	.72***	.73***						
PSF1	.52***	.53***	.64***	.59***					
PSF2	.64***	.64***	.57***	.68***	.72***				
NWF1	.67***	.69***	.63***	.51***	.43**	.39**			
NWF2	.61***	.75***	.53***	.55***	.30*	.36**	.89**		
PAST1	.53***	.56***	.61***	.54***	.62***	.49***	.65**	.60***	
PAST2	.54***	.67***	.66***	.59***	.55***	.53***	.68**	.68***	.82***

Note: All of the subtests in this table were from AIMSWeb except the PAST. LNF – Letter Name Fluency; LSF = Letter Sound Fluency; PSF = Phoneme Segmentation Fluency; NWF = Nonword Fluency; PAST = Phonological Awareness Screening Test; 1, 2 mean first administration and second administration.

*p < .05.

**p < .01

***p < .001

Table 14

Subtest Intercorrelations for Study 5

(*n* = 17)

	WID2	WID1	WA2	WA1	SPELL	EL	BW	MD	SEG
WID1	.49*								
WA2	.64**	.40							
WA1	.47	.74**	.72**						
SPELL	.66**	.60*	.62**	.59*					
EL	.47	.21	.55*	.35	.53*				
BW	.29	.69**	.32	.41	.22	.12			
MD	.47	.36	.21	.24	.51*	.44	.27		
SEG	.23	.04	.46	.25	.38	.34	.03	.22	
PAST	.60*	.78**	.53*	.57*	.59*	.44	.74**	.47	.17

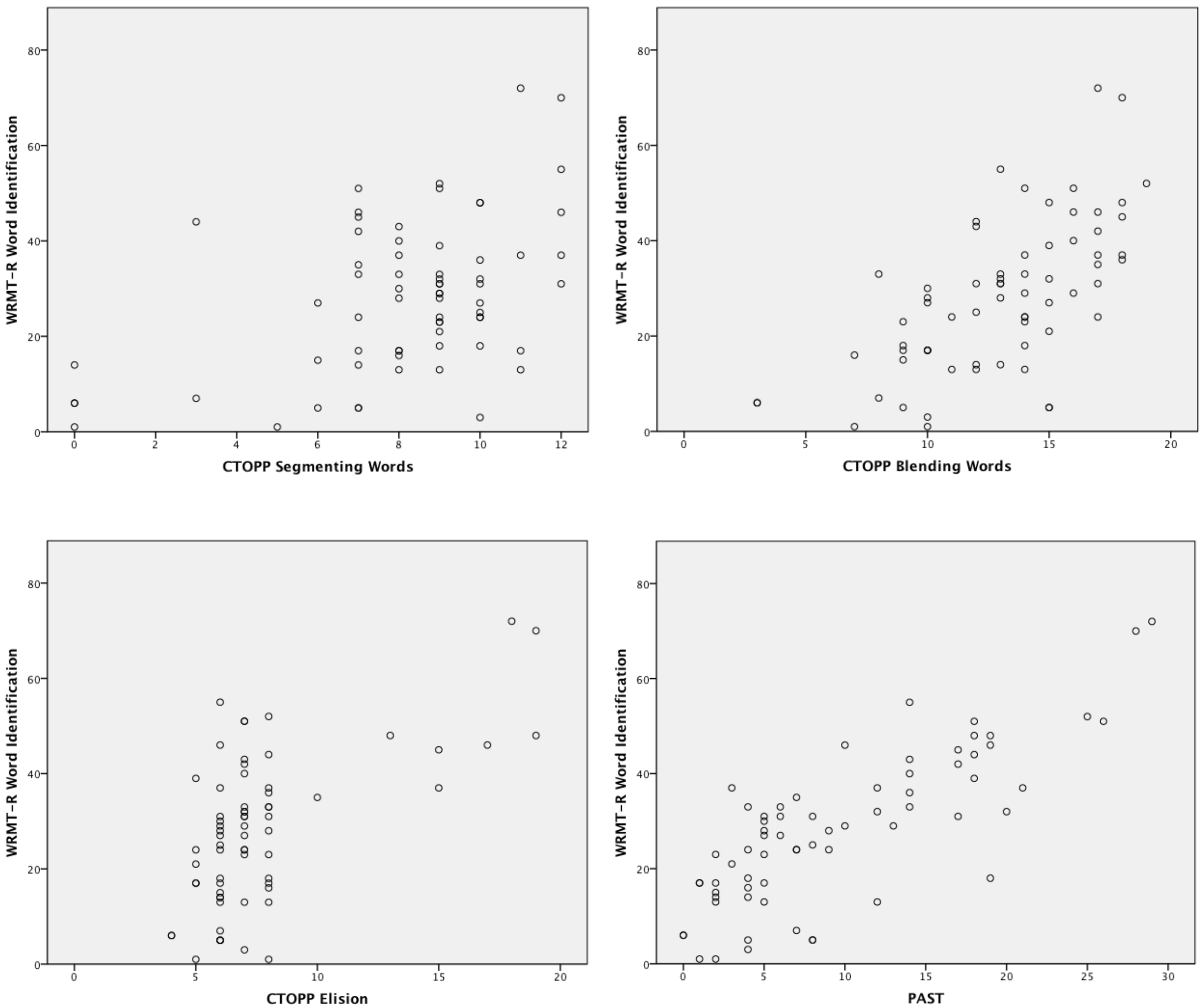
WHAT ABOUT PAST 1 VS PAST 2?

WID = WRMT-R *Word Identification*; WA = WRMT-R *Word Attack*; EWT = *Exception Words Test*; SPELL = WIAT *Spelling* subtest; EL = CTOPP *Elision*; BW = CTOPP *Blending Words*; MD = CTOPP Memory for Digits; SEG = CTOPP *Segmenting Words*; PAST = *Phonological Awareness Screening Test*;

Figure 1a

Scatter plots from Study 1 First Graders (*n* = 67)

A comparison of four phonological awareness subtests

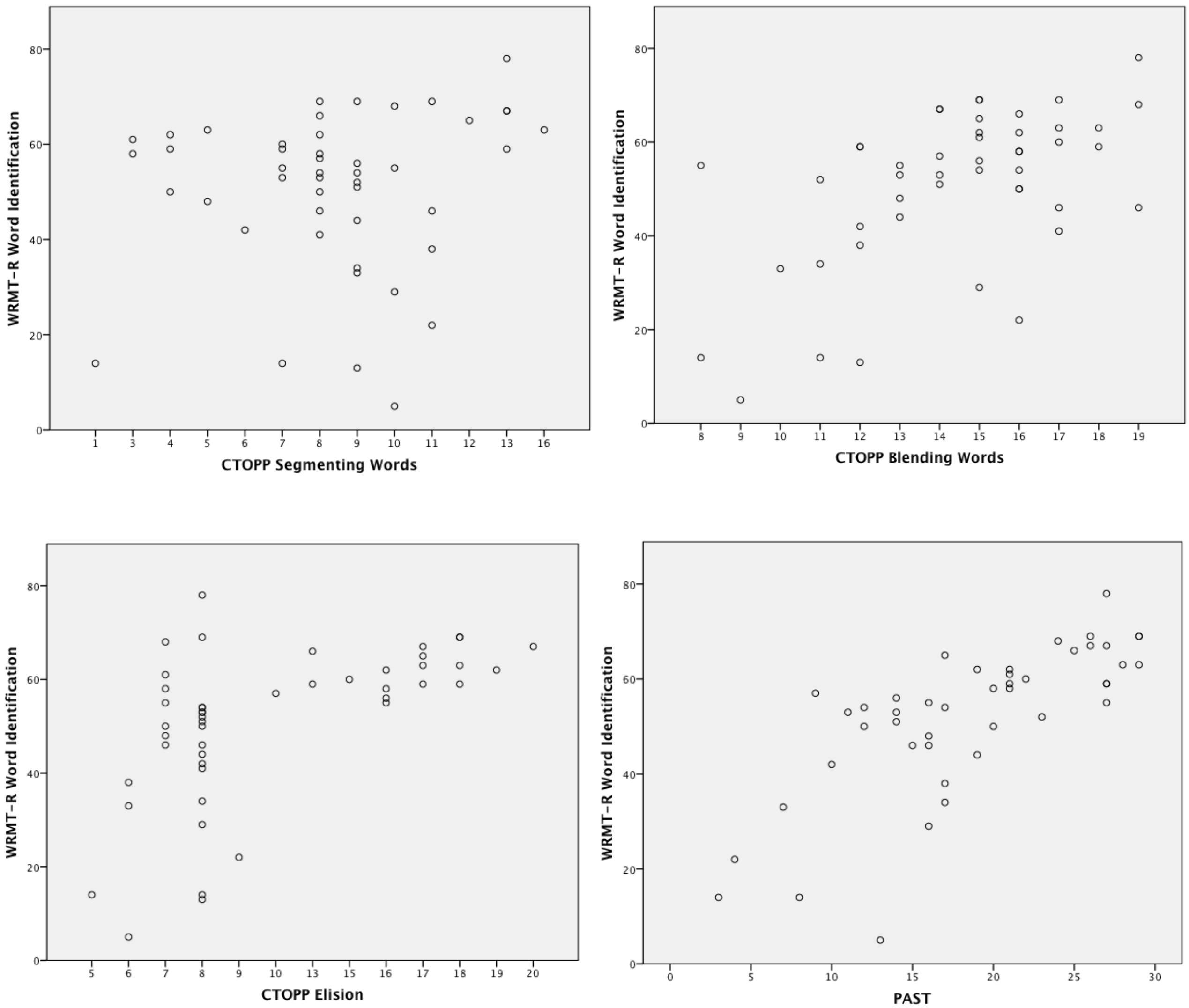


Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing*; PAST = *Phonological Awareness Screening Test*

Figure 1b

Scatter plots from Study 1 Second Graders (*n* = 50)

A comparison of four phonological awareness subtests

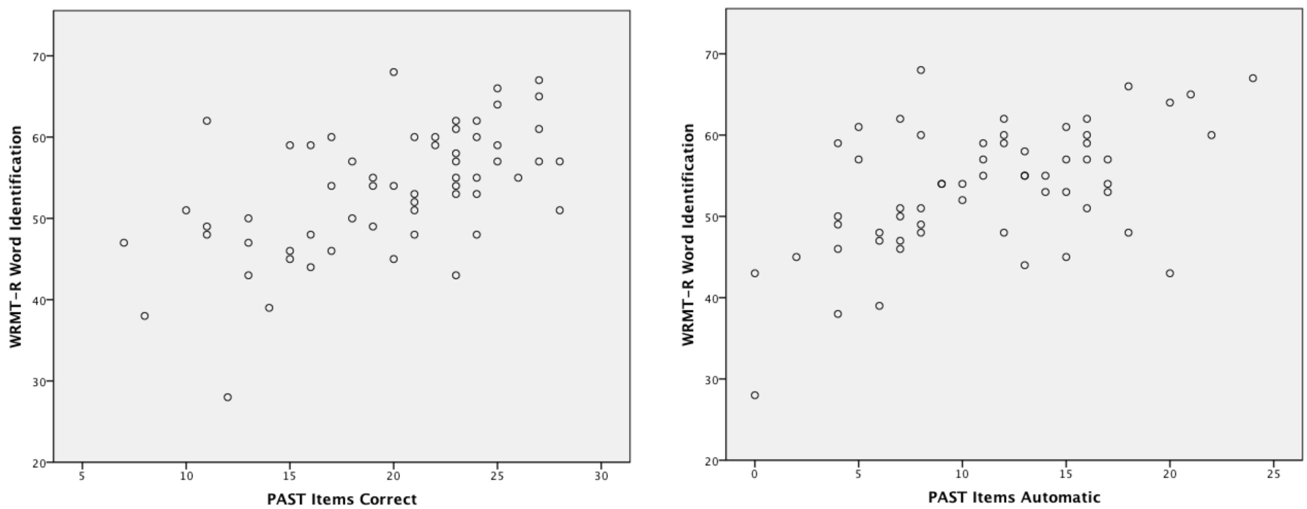


Note: WRMT-R = *Woodcock Reading Mastery Test - Revised;* CTOPP = *Comprehensive Test of Phonological Processing*; PAST = *Phonological Awareness Screening Test*

Figure 2

The PAST and Word Identification comparing items correct vs. items automatic

Grade 2 (n = 60)



Grade 5 (n = 67)