Journal of Educational Psychology
1997, Vol. 89, No. 1, 114–127

0022-0663/97/$3.00

Converging Evidence for Phonological and Surface Subtypes of Reading Disability

Keith E. Stanovich, Linda S. Siegel, and Alexandra Gottardo
University of Toronto

Using regression-based procedures introduced by A. Castles and M. Coltheart (1993), the authors identified 17 phonological and 15 surface dyslexics from a sample of 68 reading-disabled 3rd-grade children by comparing them to chronological-age (CA) controls on exception word and pseudoword reading. However, when the dyslexic subtypes were defined by reference to reading-level (RL) controls, 17 phonological dyslexics were defined but only 1 surface dyslexic. When the CA-defined subtypes were compared to RL controls, the phonological dyslexics displayed superior exception word reading but displayed deficits in pseudoword naming, phonological sensitivity, working memory, and syntactic processing. The surface dyslexics, in contrast, displayed a cognitive profile remarkably similar to that of the RL controls.

There is considerable face validity to the idea that reading disabled individuals differ among themselves in the way that they have become poor readers and in the cognitive underpinnings of their disability. Yet the field has made very little progress toward defining separable groups of disabled readers. For example, the idea that separate subgroups of poor readers could be defined based on aptitude–achievement discrepancies has not proven fruitful for the reading disabilities field (Fletcher, 1992; Fletcher et al., 1994; Shaywitz, Fletcher, Holahan, & Shaywitz, 1992; Siegel, 1988, 1989, 1992; Stanovich, 1991, 1993b, 1994; Stanovich & Siegel, 1994). The broader literature on subtyping has likewise singularly failed to converge on a common set of subtypes (see Fletcher & Morris, 1986; Lyon, 1987; McKinney, 1984; Morris & Satz, 1984; Satz, Morris, & Fletcher, 1985; Speece & Cooper, 1991; Torgesen, 1991). This older subtyping work is, in retrospect, disappointing because much of it was purely empirical and not grounded in extant theory, and some of it was grounded in theory but the theories have become dated and do not reflect the latest work in information processing or cognitive neuropsychology. In other words, the tasks used did not interface well with current models of the reading process.

There is, however, a body of subtyping work that is not subject to either of these criticisms. It is the body of work that has grown up around the study of the acquired dyslexias and the attempt to conceptualize them within the framework of current theories of word recognition. In the early 1980s, researchers (e.g., Coltheart, Masterson, Byng, Prior, & Riddoch, 1983; Temple & Marshall, 1983) began to present cases of developmental dyslexics whose performance patterns mirrored those of certain classic acquired dyslexic cases (Beauvois & Derouesne, 1979; Coltheart, Patterson, & Marshall, 1980; Marshall & Newcombe, 1973; Patterson, Marshall, & Coltheart, 1985). These cases of developmental dyslexia were interpreted within the functional cognitive architecture assumed by dual-route[1] theory (Carr & Pollatsek, 1985; Coltheart, 1978; Humphreys & Evett, 1985), which posits two routes to the lexicon—one mediated by assembled sublexical phonology and the other by a directly

Keith E. Stanovich, Linda S. Siegel, and Alexandra Gottardo, Department of Applied Psychology, Ontario Institute for Studies in Education, University of Toronto, Toronto, Ontario, Canada. Linda S. Siegel and Alexandra Gottardo are now at the Department of Educational Psychology and Special Education, University of British Columbia, Vancouver, British Columbia, Canada.

This research was supported by Grant 0GP0001607 from the Natural Sciences and Engineering Research Council of Canada.

We wish to thank Mae Burgess, Norman Himel, Margaret Lesperance, Sharon Smith, Ron Stringer, and Christine Wasson for help in data collection. We thank Jamie L. Metsala for extensive and detailed comments on an earlier version of the article and Anne Castles for kindly providing the raw data from their study.

Correspondence concerning this article should be addressed to Keith E. Stanovich, Ontario Institute for Studies in Education, University of Toronto, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6. Electronic mail may be sent via Internet to kstanovich@oise.utoronto.ca.

[1] Although current theorizing in this field has been immensely influenced by connectionist models of word recognition (Hinton & Shallice, 1991; Manis et al., 1996; Metsala & Brown, in press; Plaut et al., 1996; Plaut & Shallice, 1994; Seidenberg, 1993, 1994; Seidenberg & McClelland, 1989), we will retain dual-route nomenclature throughout this discussion. However, our adherence to this nomenclature is largely driven by explicative convenience and historical precedent rather than by a desire on our part to advance a strong position on particular micro-debates in the dual-route versus connectionist literature (see Besner, Twilley, McCann, & Seergobin, 1990; Coltheart et al., 1993; Plaut & Shallice, 1994; Seidenberg, 1993, 1994). Instead, from the standpoint of the reading disability theorist standing outside of these debates, it seems quite possible that many of these disputes arise from attempts to characterize performance at different levels of analysis (Fodor & Pylyshyn, 1988; McCloskey, 1991; Smolensky, 1988, 1989). The data patterns described remain of importance to theories of reading disability whether verbally characterized in the representational language of dual-route theory or the sub-symbolic language of connectionist theory.

addressed orthographic representation (Patterson & Coltheart, 1987).

This extrapolation from the acquired dyslexia literature to the interpretation of the performance patterns of developmental cases proved controversial, however (see Ellis, 1979, 1984; Frith, 1985; Snowling, 1983). For example, Bryant and Impey (1986) criticized the authors of the developmental case studies for not including control groups of nondisabled readers to form a context for their case descriptions. Recently, however, Castles and Coltheart (1993) have tried to answer these criticisms, first by demonstrating that their dual-route subtypes can be defined by reference to the performance of normal controls, and then by showing that the subtypes so defined are not at all rare in the reading-disabled population.

We believe that Castles and Coltheart (1993) are correct that the search for subtypes should proceed from psychological mechanisms that closely underpin the word recognition process (Baron & Strawson, 1976; Byrne, Freebody, & Gates, 1992; Murphy & Pollatsek, 1994). In this article, we show that there is reasonably convergent evidence for the subtypes defined by their procedures, however, our interpretation of the two subtypes is substantially different than theirs. First, we undertake a reanalysis of the Castles and Coltheart (1993) data based on the argument of Bryant and Impey (1986) and Manis et al. (1996) that a critical context for interpreting such subtypes is provided by a reading-level control group. We show that the patterns in the Castles and Coltheart (1993) data converge nicely with those of Manis et al. (1996). Next, we examine the Castles and Coltheart (1993) subtypes in a sample of reading-disabled children some 3 years younger than those they and Manis et al. (1996) investigated. Using ancillary tasks related to phonological and lexical processing, we examine whether the subtypes can be given a coherent conceptual interpretation.

The Castles and Coltheart (1993) Study

Castles and Coltheart (1993) analyzed the exception word and nonword reading performance of 53 dyslexic individuals and 56 nondyslexic chronological-age (CA) controls. They were motivated by a desire to distill subgroups that were relatively skilled at sublexical processing relative to lexical processing and vice versa (or, to use the terms popularized in Olson's [1994; Olson, Forsberg, & Wise, 1994] influential studies, subgroups characterized by relatively specific deficits in either phonological coding or orthographic coding). Castles and Coltheart regressed the controls' pseudoword performance on their ages and determined 90% confidence intervals around the regression line in the normal sample. They found that fully 38 dyslexics fell outside of the 90% confidence limits for this regression. In a similar vein, the dyslexics were markedly below performance expectations when naming exception words. Castles and Coltheart found that 40 of the 53 dyslexics fell below the 90% confidence intervals around the regression line based on the sample of nondyslexics.

Castles and Coltheart (1993) used these regression-based criteria to define the dissociations that would identify the phonological and surface dyslexia subtypes in their sample. They found that 18 of the 53 participants fulfilled one or the other of the dissociation criteria. Specifically, 10 participants fit the surface dyslexia pattern: They were below the normal range in exception word reading but were within the normal range on pseudoword reading. Eight participants fit the phonological dyslexia pattern: They were below the normal range in pseudoword reading but were within the normal range on exception reading. These 18 cases might be termed the "hard" subtypes. They fit the classic dissociation definitions: normal on one subprocess (at least by these particular operational criteria) and subnormal on the other.

However, Castles and Coltheart (1993) argued that additional cases of subtypes could be identified, not on the basis of abnormal performance on one measure and normal performance on the other, but on *relative* imbalances on the two tasks among children who might well show depressed performance on both (we might term cases defined in this way the "soft" subtypes). Castles and Coltheart argued that a principled way of operationalizing this imbalance was, again, by reference to the performance of the normal control group.

The soft subtypes were defined by running a regression line with 90% confidence intervals through the Exception Word × Pseudoword plot for the control children. This regression line and confidence intervals were then superimposed on the scatterplot of the performance of the dyslexic sample. Participants falling below the lower confidence interval in this plot and *not* its converse qualify for the *soft surface dyslexia* subtype: They are unusually impaired on exception word reading relative to their performance on pseudowords. An analogous regression defines the *soft phonological dyslexia* subtype. Using this procedure, Castles and Coltheart (1993) defined 16 surface dyslexics and 29 phonological dyslexics. They thus argued that the vast majority of dyslexics in their sample (45 out of 53) displayed some type of dissociation and they concluded that "a clear double dissociation exists between surface and phonological dyslexic reading patterns. . . . [I]t would seem that these reading patterns are not rare phenomena, but are quite prevalent in the developmental dyslexic population" (p. 174).

Reanalysis of the Castles and Coltheart (1993) Data

Nevertheless, conceptual and statistical interpretation of the Castles and Coltheart (1993) data is problematic for a reason argued by Bryant and Impey (1986) over 10 years ago: the lack of reading-level controls. The problem is that if the processing tradeoffs involving the lexical and sublexical procedures are specifically linked with the overall level of word recognition that the reader has attained, then using the generally higher performance relationships of the CA group as the benchmark might represent an inappropriate extrapolation. A reanalysis of the Castles and Coltheart data[2] serves to confirm these fears. When the performance of the CA controls on exception words is plotted against reading age and the performance on pseudowords is plotted against reading age, both variables display statistically significant quadratic trends ($p < .001$, in the case of pseudowords). In precisely the range of reading ages where the reading-disabled sample resides, the slope relating performance to reading age is steeper.

Another way of viewing this problem is to note that when the entire sample is considered, the slope of the function relating exception word performance to reading age (in months) is steeper for the reading-disabled sample than for the CA controls (.260 vs. .128). However, when the range of reading ages is restricted to the lower range (<117 months) where there is overlap between the CA controls and reading-disabled sample, there is no difference in slopes (.260 vs. .310). In a similar way, when the entire sample is considered, the slope of the function relating pseudoword performance to reading age (in months) is steeper for the reading-disabled sample than for the CA controls (.305 vs. .114). However, when the range of reading ages is restricted to the lower range (<117 months) where there is overlap between the CA controls and reading-disabled sample, there is no difference in slopes (.305 vs. .378).

In short, the difference in the growth functions when the entire sample of CA controls is compared with the sample of dyslexics is simply a function of the differing distributions of the two samples across the reading-age continuum. The steeper slope displayed by the dyslexics is not a function of being dyslexic—it is simply a property of these particular pseudowords and exception words being given to children of these particular reading levels. When reading at the same level, control children display exactly the same slope (this, of course, is a variant of the arguments for reading-level controls that have appeared in the literature before; see Bryant & Goswami, 1986; Bryant & Impey, 1986). In fact, one can easily see that a regression line dominated by high reading-age control children is an inappropriate one for the dyslexic children by simply pondering the fact that it is equally inappropriate for *normal* children of reading ages similar to the dyslexics.

The moral here is that if the processing tradeoffs involving the lexical and sublexical procedures are specifically bound up with the overall level that the reader has attained, then extrapolating from the reading patterns of children at a higher reading level is an inappropriate way of defining abnormal patterns of processing skills at a lower reading level. This admonition applies equally to the procedures used to define the soft subtypes as it does the hard subtypes. (A comparison of Figures 5 and 6 in the Castles and Coltheart (1993) article reveals that the bivariate distribution of the performance of the dyslexic participants is virtually outside of the bivariate space of the CA controls.)

In a recent study, Manis et al. (1996) added an important context for the Castles and Coltheart type of analysis: a reading-level control sample. First, they replicated the Castles and Coltheart (1993) procedures with chronological-age controls, as in the original investigation. Out of their sample of 51 reading-disabled students, they found only half as many hard subtype cases as did Castles and Coltheart: 5 surface dyslexics and 5 phonological dyslexics. (They found, as did Castles and Coltheart, that most dyslexics were significantly depressed on their use of *both* the sublexical and lexical procedures. In our search for subtypes, we should not lose sight of the fact that most dyslexics are low on both.) However, they found, as did Castles and Coltheart, that large numbers of soft subtypes could be defined by using the exception word–pseudoword regression criteria derived from CA controls: 17 participants were soft phonological dyslexics and 15 were soft surface dyslexics.

Manis et al. (1996) next did a soft subtype analysis using a regression line based on the performance of reading-level controls rather than CA controls. Twelve of the 17 phonological dyslexics also qualified for that subtype on the basis of the reading-level (RL) analysis. In contrast, 5 children defined as phonological dyslexics on the basis of the CA analysis were no longer so when an RL control group was used. However, an even more striking outcome was obtained when the performance of the surface dyslexics was examined. Only 1 out of 15 surface dyslexics qualified for this subtype label when an RL control group was used. As a result of these findings, Manis et al. (1996) concluded that "the phonological dyslexic profile represents a specific deficit in phonological processing, whereas the surface dyslexic profile represents a more general delay in word recognition" (p. 179).

In fact, the finding of Manis et al. (1996) actually converges with data patterns evident in the Castles and Coltheart (1993) study itself. Given that the latter study was a CA match investigation, one might wonder how this was possible. It is possible because the Castles and Coltheart study shares a characteristic of many other studies in the reading disabilities literature. That characteristic is that even when participants are reported to be matched at a particular chronological age, reading age, IQ, or whatever other variable, there is often enormous variability around the point of the match (as identified by a statistic such as the mean). This is certainly true in the Castles and Coltheart study. The CAs in their study spanned 7½ years (90 months to 179 months) and the reading ages spanned 7 years (78 months to 163 months). It is important to note, however, at the lower reading ages there is overlap between the groups. We thus formed matched reading-level groups of 17 nondyslexic children and 40 dyslexic children from the Castles and Coltheart data, a comparison not examined in their original article. Table 1 indicates that although the match was less than perfect (there was an almost 3 months difference in reading age), the difference was not statistically significant. The reading-level controls outperformed the dyslexics on all three stimulus types. The difference was much larger on pseudowords, and the interaction between stimulus type and participant group was highly significant ($p < .001$). The dyslexics named about the same number of pseudowords as exception words, whereas the RL controls named about six

[2] We thank Anne Castles for providing the raw data from their study that allowed us to conduct a series of replots and reanalyses.

Table 1
Mean Differences Between the Reading-Level Matched Dyslexics (N = 40) and the Nondyslexics (N = 17) in the Castles and Coltheart (1993) Study

Variable	Dyslexics	Controls	*t* value
Reading age	101.0	103.9	1.34
Exception words	13.0	14.9	2.17*
Pseudowords	13.8	20.6	3.61***
Regular words	22.5	26.0	2.80**

$* p < .05$. $** p < .01$. $*** p < .001$, all two-tailed.

more pseudowords than exception words—a very large difference. Thus, the reading-level match from the Castles–Coltheart data replicates the classic finding of a dyslexic pseudoword reading deficit in an RL match (Rack, Snowling, & Olson, 1992; Stanovich & Siegel, 1994).

Figure 1 illustrates one of the plots that identifies the soft subtypes; however, this time we used only the RL control group as a benchmark rather than the full CA-matched control group. The number of pseudowords read correctly is plotted against the number of exception words read correctly. The performance of the reading-disabled children is represented by squares and the performance of the nondisabled RL-matched group is represented by triangles. The regression line and confidence intervals displayed in the figure are based on the data from the 17 RL controls (the triangles). This plot in part identifies the soft phonological subtype (children low on pseudoword reading relative to exception word reading). There are 15 phonological dyslexics according to this criterion.

Figure 2 displays the performance of the dyslexics plotted so as to identify surface dyslexics (children low on exception word reading relative to pseudoword reading). The number of exception words read correctly is plotted against the number of pseudowords read correctly. The performance of the reading-disabled children is again represented by squares and the performance of the nondisabled RL-matched group is represented by triangles. The regression line and confidence intervals displayed in the Figure are again based on the data from the 17 RL controls (the triangles). Figure 2 indicates that the Castles and Coltheart (1993) data patterns themselves converge with the findings of Manis et al. (1996). Most surface dyslexics disappear when a reading-age control is used, however only 2 are left in the Castles–Coltheart sample (one just barely making the cutoff). Thus, a reanalysis of the original Castles and Coltheart data replicates the trend demonstrated by Manis et al. (1996). When an RL control group is used, surface dyslexics defined by a CA-match are disproportionately eliminated.

Phonological and Surface Subtypes in a Younger Sample

We have conducted a subtype analysis of the Castles and Coltheart type on a sample of children who were considerably younger than those studied by Manis et al. (1996) and in the post hoc analysis of the Castles and Coltheart (1993) study. Our study extended beyond their findings in three ways. As Table 2 indicates, our dyslexics and RL controls were considerably younger than the children in the other studies. Thus, we examined whether the results generalize to earlier reading levels and how early the subtypes can be reliably identified. Second, both the Castles and Coltheart and the Manis et al. studies examined samples that varied widely in age. In contrast, our dyslexics, as well as their CA controls, were all third graders and our RL controls were all first and second graders. Finally, in our battery, unlike in the Castles and Coltheart study, we included a variety of other tasks that could provide some converging validation for the subtypes (see also Manis et al., 1996).

Method

Participants

The reading-disabled sample consisted of 68 third-grade children (29 boys and 39 girls, mean age = 107.5 months) attending suburban schools in a large metropolitan area. These children all scored below the 25th percentile on the reading subtest of the Wide Range Achievement Test—Revised (WRAT-R; Jastak & Wilkinson, 1984). The mean percentile rank of the less-skilled group on this subtest was 10.6. The chronological-age controls were 44 children (16 boys and 28 girls, mean age = 107.8 months) recruited from the same schools who scored above the 30th percen-

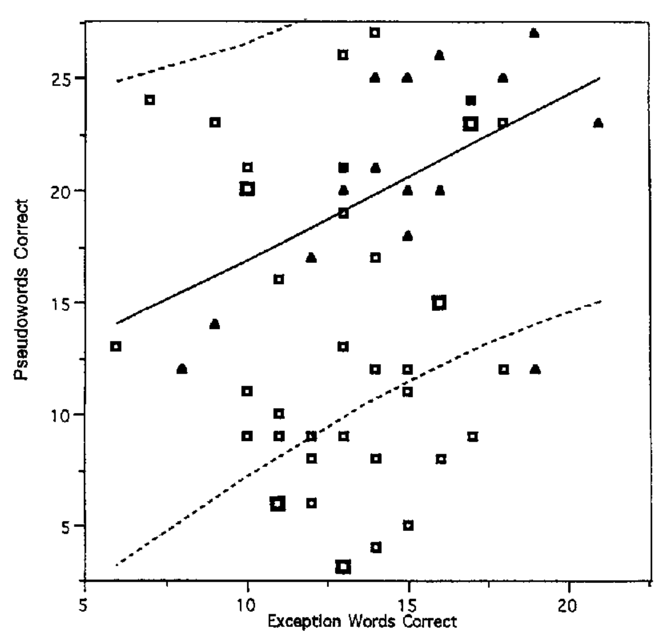

Figure 1. Performance on pseudoword reading plotted against exception word reading for the reading-disabled children (squares) and reading-level (RL) controls (triangles) in the Castles and Coltheart (1993) data. The regression line and confidence intervals were derived from the data of the RL controls. Larger squares indicate two individuals with reading disability with the same scores, and filled squares indicate that two individuals, one from each group, have the same scores.

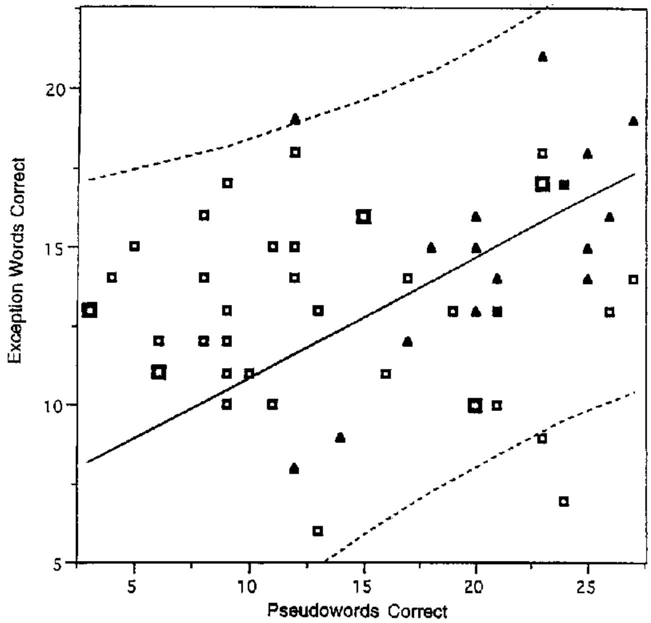

Figure 2. Performance on exception word reading plotted against pseudoword reading for the reading-disabled children (squares) and reading-level (RL) controls (triangles) in the Castles and Coltheart (1993) data. The regression line and confidence intervals were derived from the data of the RL controls. Larger squares indicate two individuals with reading disability with the same scores, and filled squares indicate that two individuals, one from each group, have the same scores.

tile on the reading subtest of the WRAT–R. The mean percentile rank of the chronological-age controls on this subtest was 58.5.

These children represented a subsample derived from testing approximately 200 children: all of the third graders in three schools within the same school district. The district serves a large multicultural population, and average achievement levels in the schools tested are substantially below the norms of most standardized tests. Thus, our relatively liberal criterion for reading disability combined with the requirement that the controls achieve above the 30th percentile resulted in an overrepresentation of reading-disabled children and an underrepresentation of controls. We have examined a variety of more stringent criteria for reading disability and more liberal criteria for control status, but these different (and equally arbitrary) cutoff points do not materially affect any of the data patterns. At the time of the study, educational personnel felt that many of the lower socioeconomic status (SES) students were not well served by the lack of emphasis on alphabetic coding in the curriculum (see Adams & Bruck, 1993; Hatcher, Hulme, & Ellis, 1994; Iversen & Tunmer, 1993; Share & Stanovich, 1995; Vellutino, 1991). We eliminated all children who were not native English speakers (36.9% of the sample) or had reported histories of speech, language, or hearing difficulties (9.2% of the sample).

The reading-level controls consisted of 23 first- and second-grade children (13 boys and 10 girls, mean age = 88.9 months) whose mean raw score on the WRAT–R was matched to that of the less-skilled third graders (see Table 3). The mean percentile rank of the reading-level controls on this subtest was 49.5. The children were tested in May and June of the school year.

Tasks

Standardized measures. In addition to the reading subtest of the WRAT–R, the children were administered the spelling and arithmetic subtests of the WRAT–R and the Form G of the Word Attack subtest of the Woodcock Reading Mastery Tests (Woodcock, 1987), a test of pseudoword reading. As an additional measure of word recognition ability, the children were administered the Word Identification subtest of the Woodcock Reading Mastery Tests (Form G).

Experimental words and pseudowords. The two reading disability subtypes were defined by their performance on a set of experimental exception words and pseudowords largely drawn from the work of Coltheart and Leahy (1992) and from other studies in which word recognition mechanisms have been studied (e.g., Laxon, Smith, & Masterson, 1995; Patterson & Morton, 1985; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Waters & Seidenberg, 1985). There were 21 exception words (bear, wood, pint, sweat, broad, steak, worse, shove, shoe, debt, yacht, ghost, island, ocean, doubt, rhyme, aisle, sword, muscle, amoeba, receipt) that represented a mixture of irregular or inconsistent words (see Coltheart & Leahy, 1992) having vowel pronunciations that are of low frequency based on small-unit, context-free counts (see Berndt, Reggia, & Mitchum, 1987) and also having orthographic neighbors that conflict with their pronunciation (e.g., pint, steak), and also so-called "strange" or "unique" words (Seidenberg et al., 1984; Waters & Seidenberg, 1985), which are not only pronounced irregularly but which have unusual orthographic patterns and few orthographic neighbors (e.g., yacht, aisle, amoeba).

The children also attempted to pronounce 30 pseudowords (ving, fump, drace, biss, pask, hane, drack, lail, fide, hile, stell, vind, tralf, pold, grall, bould, nalk, jook, fralt, rild, stull, vood, sost, nush, bove, trome, zove, fown, slear, yone). To be scored as correct, the vowel in the pseudoword had to be pronounced in accordance with some real words with the identical rime unit. For example, the pseudoword *bove* would have to be pronounced to rhyme with one of the real words *cove, love,* and *move.* Finally, the children attempted to pronounce 18 regular words (glide, stiff, press, blame, brace, tribe, smoke, crane, puff, bead, cove, paid, phone, dome, slave, dive, speak, hull) that represented a mixture of regular or consistent words (e.g., stiff, press, blame) containing vowels pronounced with their most frequent small-unit correspondence (Berndt et al., 1987) and with word bodies (iff, ess, ame) all having consistent pronunciations, and also regular inconsistent words (e.g., paid, slave, dive) containing vowels pronounced with their most frequent vowel pronunciation but which have neighbors whose shared word bodies are pronounced differently.

Phonological sensitivity. Phonological sensitivity was assessed by administering Rosner's Auditory Analysis Test (Rosner & Simon, 1971), which involves both syllable and phoneme deletion. The child was told that "We are going to play a word game. I am going to say a word and I want you to say it the same way I do. Then I am going to tell you what part to take off and then I want

Table 2
Mean Age Differences (in Months) Among the Three Studies

Variable	Dyslexics	Controls
Post hoc analysis of Castles and Coltheart (1993)	137.9	102.0
Manis et al. (1996)	149.2	102.0
Present investigation	107.5	88.9

Table 3
Mean Differences Between the Dyslexics (N = 68) and the Reading-Level Controls (N = 23)

Variable	Dyslexics	Controls	*t* value
WRAT–R Reading[a]	51.2	51.0	0.12
WRAT–R Spelling[a]	33.1	31.4	1.86
Woodcock Word Identification[a]	46.8	42.6	1.83
WRAT–R Arithmetic[a]	23.5	20.3	5.72**
Woodcock Word Attack[a]	12.1	15.5	−2.25*
Rosner AAT	16.8	21.1	−2.16*
Wordlikeness choice	11.8	11.8	0.01
Exception words	6.9	6.4	0.52
Regular words	8.4	9.2	−0.73
Pseudowords	13.9	16.8	−2.05*
Syntactic processing[b]	−.148	.235	−2.26*
Working memory[b]	−.134	.310	−2.77*

Note. WRAT–R = Wide Range Achievement Test—Revised; Rosner AAT = Rosner Auditory Analysis Test.
[a] raw score. [b] z score.
* $p < .05$. ** $p < .001$, all two-tailed.

you to say what's left." The child was then given two practice items followed by the 40 items of the test. Participants were asked to delete syllables, single phonemes from initial and final positions in words, and single phonemes from blends. A maximum of one repetition was allowed per item if the child requested a repetition or had not responded within approximately 15 s. The 40 items were arranged in approximate order of difficulty and testing was discontinued after five consecutive error responses. The maximum score on the task was 40.

Wordlikeness choice. In this task, the children were shown two nonword strings (e.g., filv–filk, lund–dlun). They were told that neither string looks or sounds like an actual word but that one letter string is more *like* a word. One member of each pair contained an orthographic sequence that either never or rarely occurs in English in that particular position in a word (e.g., filv, dlun). The subject's score is the number of times that the nonword without the illegal or low-frequency letter string was chosen. Although this task undoubtedly implicates phonological coding to some extent, the coding of frequent and infrequent orthographic sequences in memory should be a substantial contributor to performance. There were 17 trials and the raw number correct was used in the analyses that follow.

Syntactic processing. Syntactic processing skills were measured by using a sentence judgment task and a sentence correction task. The same stimuli were used for both the judgment and correction tasks. The syntactic judgment task was administered at the beginning of each individual testing session and the syntactic correction task was presented at the end of each individual testing session. Thus, the two syntactic tasks were separated maximally in terms of time, so that the child's performance on the judgment task would have minimal impact on performance on the correction task. In the syntactic judgment task, the participant was asked whether orally presented sentences are correct or incorrect. There were 25 sentences with errors presented and 10 foils (sentences without errors). The maximum score on the task was thus 35. The error sentences included errors in clause order within sentences, errors in word order within clauses, errors in subject–verb agreement, errors in subject–copula verb agreement, and errors in function word usage.

The same 25 error sentences that were used for the judgment task were orally presented for the sentence correction task in which the children were asked to "fix" the sentence. The scores on the judgment and correction tasks were combined into a single syntactic processing score by first converting the scores on each task to z scores and then averaging these z scores.

Verbal working memory. Verbal working memory was measured using a variation of a memory task developed by Daneman and Carpenter (1980). Children were required to listen to a series of statements and to respond *true* or *false* (or *yes–no*). After responding to each of the sentences in a set, the child was required to recall the final word of each sentence in the set. The children received 3 two-item sets, 3 three-item sets, and 3 four-item stimulus sets. Across the 9 sets of items, the children responded to 27 true–false sentences and attempted to recall 27 words. The mean recall score was 14.5 items (SD = 4.0) out of a total of 27. The mean number of true–false items answered correctly was 21.9 (SD = 2.7). A composite variable that combined the recall and true–false performance was used in the analyses presented later. The composite measure was formed by standardizing the scores on the recall and true–false parts of the verbal working memory task and then averaging these standard scores. Exactly the same trends were apparent when only the recall scores were used.

Procedure

All of the tasks were administered in a single individual session and a single group session that included several other tasks not reported here. In the individual session, to maximally separate the two syntactic tasks the syntactic judgment task was administered at the beginning of the individual testing session and the syntactic correction task was administered at the end of the session. The remaining tasks were presented in order: WRAT–R Reading subtest, Rosner Auditory Analysis Test, experimental pseudowords, experimental regular and exception words, verbal working memory task, Woodcock Word Attack, Woodcock Word Identification, and wordlikeness choice (with occasional deviations from this order necessitated by material sharing). The WRAT–R Spelling subtest and WRAT–R Arithmetic subtest were administered in a group session.

Results

Recall that the soft dyslexia subtypes were defined by plotting pseudoword performance against exception word performance (and vice versa) and examining the 90% confidence intervals around the regression line determined from the CA control group. A phonological dyslexic is a child who is an outlier when pseudowords are plotted against exception words but is within the normal range when exception words are plotted against pseudowords. Surface dyslexics are defined conversely. Figure 3 displays the data from our 68 third-grade reading-disabled children and plots experimental exception word performance against experimental pseudoword performance. The regression line and confidence intervals from the 44 CA controls in our sample are also displayed. All four groups that are defined by conjoining the results of this with the converse plot not shown (pseudoword performance against exception word performance) are indicated. Specifically, the points labeled with Ys are the surface dyslexics (low in the Exception Word × Pseudoword plot and in the normal range on the converse plot), the triangles are the phonological dyslexics

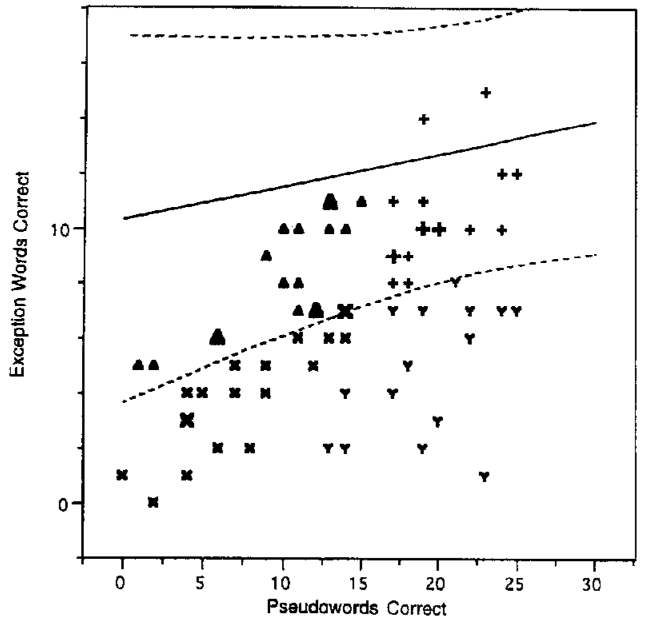

Figure 3. Performance on exception word reading plotted against pseudoword reading for the reading-disabled children in this study. The regression line and confidence intervals were derived from the data of the chronological controls. Y = surface dyslexics, ▲ = phonological dyslexics, × = low on both, + = low on neither. Larger points indicate two individuals with the same scores.

(low in the Pseudoword × Exception Word plot and in the normal range on the converse plot), the Xs are participants who are low on both measures, and the crosses represent individuals who are low on neither.

One interesting difference between our results and those involving the CA controls in the Castles and Coltheart (1993) and Manis et al. (1996) investigations concerns the fact that Manis et al. found that only 9.8% of their sample were outside the regression criterion on both measures, and Castles and Coltheart found only 5.7% of their sample low on both. In contrast, in our younger sample, 27.9% of the dyslexics (19 out of 68 children) were low on both types of stimuli. Perhaps these findings indicate that, with development, there is increasing dissociation between lexical and sublexical processes in dyslexics. The proportion of surface dyslexics was fairly similar across the three studies (30.2% in the Castles & Coltheart study, 29.4% in Manis et al., and 22.1% in our sample). In contrast, the proportion of phonological dyslexics in the Castles and Coltheart study (54.7%) was higher than that observed in the other two investigations (33.3% in Manis et al., and 25.0% in our sample).[3]

Table 3 displays the comparisons between the 23 reading-level controls and the 68 dyslexics in our sample. The groups were matched closely on their WRAT reading raw scores. The dyslexics scored somewhat higher on the WRAT spelling subtest and on the Woodcock Word Identification subtest, perhaps indicating some degree of regression in the matched groups. The sample can be more closely matched on these variables at a cost in sample size, but it does not materially affect the results. The older dyslexics were superior in arithmetic performance, a common finding in an RL match (see Stanovich & Siegel, 1994). On two measures of pseudoword reading (the Woodcock Word Attack and the experimental pseudowords), we replicated the finding of a dyslexic deficit in an RL match (Rack et al., 1992; Stanovich & Siegel, 1994). On a measure of phonological sensitivity (the Rosner Auditory Analysis Test), the dyslexics displayed a significant deficit, consistent with other research in the literature (Bowey, Cain, & Ryan, 1992; Bradley & Bryant, 1978; Bruck, 1992; Bruck & Treiman, 1990; Olson, Wise, Conners, & Rack, 1990).

No differences were displayed on the set of exception and the set of regular words. On a measure of orthographic processing (a wordlikeness choice task), the two groups displayed no difference. The latter task has sometimes revealed superior processing for reading-disabled individuals in an RL match (Siegel, Share, & Geva, 1995) and sometimes not (Stanovich & Siegel, 1994). Our results are, however, consistent with all previous work with this task in indicating that reading-disabled individuals do not show an RL deficit and that their orthographic processing problems are less severe than their phonological processing problems (Olson, Wise, Conners, Rack, & Fulker, 1989; Siegel et al., 1995; Stanovich & Siegel, 1994). Finally, the reading-disabled children performed significantly worse than the younger controls on the measures of syntactic processing skill and verbal working memory, perhaps indicating that these tasks are in part tapping their core phonological deficit, for which there is substantial evidence (Bruck, 1992; Goswami & Bryant, 1990; McBride-Chang, 1995a, 1995b; Olson, 1994; Perfetti, 1994; Shankweiler, Crain, Brady, & Macaruso, 1992; Shankweiler et al., 1995; Share, 1995; Share & Stanovich, 1995; Siegel & Ryan, 1988; Stanovich, 1988, 1991; Stanovich & Siegel, 1994; Wagner, Torgesen, & Rashotte, 1994).

Figure 4 displays the performance of the dyslexics plotted so as to identify phonological dyslexics (children low on pseudoword reading relative to exception word reading). The number of pseudowords read correctly is plotted against the number of exception words read correctly. The performance of the reading-disabled children is represented by Xs and the performance of the nondisabled RL-matched group is represented by squares. The regression line and confidence intervals displayed in the figure are based on the data from the 23 RL controls (the squares in Figure 4).

Figure 5 shows the performance plotted so as to identify surface dyslexics (children low on exception word reading relative to pseudoword reading). The number of exception words read correctly is plotted against the number of pseudowords read correctly. The performance of the reading-disabled children is again represented by Xs and the performance of the nondisabled RL-matched group is represented by squares. The regression line and confidence

[3] A 90% confidence interval was used in the present study and in the Castles and Coltheart (1993) study. Manis et al. (1996) used a 95% confidence interval.

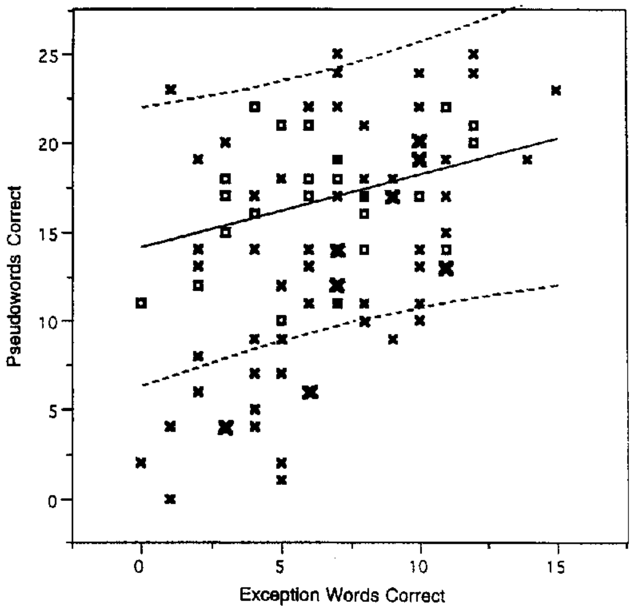

Figure 4. Performance on pseudoword reading plotted against exception word reading for the reading-disabled children (Xs) and reading-level (RL) controls (squares) in this study. The regression line and confidence intervals were derived from the data of the RL controls. Larger Xs indicate two individuals with reading disability with the same scores, and filled squares indicate that two individuals, one from each group, have the same scores.

intervals displayed in the figure are again based on the data from the 23 RL controls (the squares).

Seventeen children were identified as phonological dyslexics by using these two scatterplots and regressions. (Not all of these were identical to those identified from the CA regression lines. Due to differing slopes and intercepts, some CA phonological dyslexics were low on neither measure in the RL analysis, and some children who were low on both in the CA analysis were phonological dyslexics in the RL analysis.) Figure 5 indicates that, consistent with the findings from the older sample of Manis et al. (1996) and our reanalysis of the Castles–Coltheart data presented earlier, surface dyslexics virtually disappear when a reading-age control is used: Only 1 is left in our sample. This is consistent with the 2 found in the Castles–Coltheart sample and 1 found in the Manis et al. sample. In contrast, all three samples identify substantial numbers of phonological dyslexics in RL-control comparisons.

The results of all of these analyses suggest that the surface dyslexics defined by CA comparisons appear to be children with a type of reading disability that could be characterized as a developmental lag. The performance of surface dyslexics is in no way unusual, at least in comparison to other normal readers at the same level of overall ability (see Beech & Harding, 1984; Stanovich, Nathan, & Zolman, 1988; Stanovich & Siegel, 1994). In contrast, phonological dyslexia defined by comparison with a CA control group seems to reflect true developmental deviance. This conclusion is reinforced by examining performance comparisons between the surface dyslexics and RL controls on the other variables contained in our performance battery (see also Manis et al., 1996). Table 4 presents these comparisons. It is apparent that on only one variable (WRAT Spelling) were the two groups significantly different. The two groups of children performed similarly on several tasks not used to define the dyslexic subtypes (Rosner AAT, wordlikeness choice task, and two subtests of the Woodcock) as well as measures of syntactic processing and verbal working memory that were included in this study. The latter two measures add to the picture of developmental lag that seems to characterize the surface subtype: These children had syntactic processing skills and verbal memory skills commensurate with their reading-level controls.

Comparisons of the phonological dyslexics with the RL controls are in marked contrast to those involving the surface dyslexics. Table 5 indicates that here, there were several significant differences between the groups. The phonological dyslexics were markedly inferior on not only the experimental pseudowords that in part defined the groups but also on the Woodcock Word Attack subtest (not used to define the groups). Their phonological problems were further indicated by a significant deficit in phonological sensitivity as indicated by their performance on the Rosner Auditory Analysis Test. They were significantly better at reading exception words. One very interesting finding that serves to confirm the developmental deviancy of this group in the phonological–language domain was that phonological dyslexics performed significantly worse than these younger controls on the measures of syntactic processing skill and

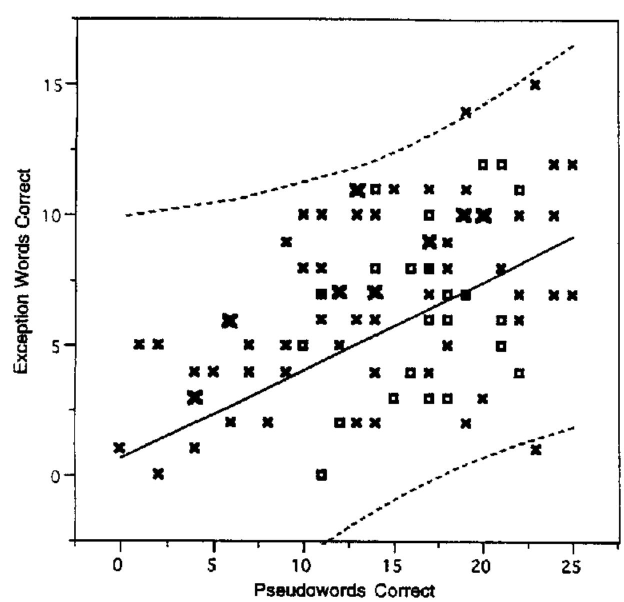

Figure 5. Performance on exception word reading plotted against pseudoword reading for the reading-disabled children (Xs) and reading-level (RL) controls (squares) in this study. The regression line and confidence intervals were derived from the data of the RL controls. Larger Xs indicate two individuals with reading disability with the same scores, and filled squares indicate that two individuals, one from each group, have the same scores.

Table 4
Mean Differences Between the Surface Dyslexics (N = 15) and the Reading-Level Controls (N = 23)

Variable	Dyslexics	Controls	*t* value
WRAT–R Reading[a]	53.1	51.0	0.96
WRAT–R Spelling[a]	33.7	31.4	2.09*
Woodcock Word Identification[a]	46.4	42.6	1.09
Woodcock Word Attack[a]	15.4	15.5	−0.04
Exception words	4.8	6.4	−1.66
Regular words	11.3	9.2	1.45
Pseudowords	19.2	16.8	1.96
Rosner AAT	21.5	21.1	0.17
Wordlikeness choice	11.2	11.8	−0.69
Syntactic processing[b]	.256	.235	0.12
Working memory[b]	.090	.310	−1.03

Note. WRAT–R = Wide Range Achievement Test—Revised; Rosner AAT = Rosner Auditory Analysis Test.
[a] raw score. [b] *z* score.
$*p < .05$.

verbal working memory, probably indicating that their core phonological deficit impairs performance on these tasks.

We explored the stability and reliability of these group classifications by using alternative measures to classify the children. For example, in one alternative classification, we used the Woodcock Word Attack subtest as an indicator of sublexical processing and the Word Identification subtest as an indicator of lexical processing. The same regression-based procedure as before was used with the CA controls as a baseline but with these two new criterion variables. In another alternative classification, we used the Woodcock Word Attack subtest as an indicator of sublexical processing and the Spelling subtest of the WRAT as an indicator of lexical processing (see Hanley, Hastie, & Kay, 1992; Juel, Griffith, & Gough, 1986). There was, in fact, no child who fulfilled the criteria for surface dyslexia across all three of the classifications carried out on the data (the original classification and these two new ones). In contrast, 6 children were classified as phonological dyslexics in all three analyses. This finding converges with the way the two subtypes profiled in terms of developmental deviance versus delay. The phonological subtype appears to be a much more distinct subtype. These individuals are developmentally deviant and a subset of these children show phonological and orthographic processing capabilities that remain dissociated regardless of changes in how those two hypothesized processes are assessed. In contrast, surface dyslexia appears to be a developmental delay, and it is much less robust in the sense that individuals with this type of dissociation are not always consistently defined across changes in methods of measuring lexical and sublexical processing in a CA match.

In thinking about subtypes, it is also important not to ignore the "deviant on both" group—the children below the CA control group confidence intervals for both pseudowords and exception words. As noted previously, this group was much larger in our sample of younger participants (27.9% of the dyslexics) than in the Manis et al. (1996) and Castles–Coltheart (1993) samples who were 2½ to 3½ years older. We conjecture that this "deviant on both" group are perhaps phonological dyslexics of the future, a hypothesis supported by the results displayed in Table 6, which compares the performance of the phonological dyslexics to that of the deviant on both group. Here we see that the both-deviant group shares *all* of the phonological deficits of the phonological dyslexics: They are equally impaired at reading pseudowords and in phonological sensitivity. They share the syntactic processing problems and verbal working memory deficits—deficits that may well arise from processing problems at the phonological level (see Gottardo, Stanovich, & Siegel, 1996; Shankweiler et al., 1992, 1995). The differences between the groups arise because the phonological dyslexics are better at reading words, particularly exception words.

Discussion

Given that IQ-based aptitude–achievement discrepancies have been shown to have low returns for the reading disabilities field as mechanisms for demarcating conceptually interesting subtypes (e.g., Stanovich & Siegel, 1994), it appears that Castles and Coltheart (1993) are correct that the search for subtypes should proceed from psychological mechanisms that closely underpin the word recognition process. In our study, we have explored the implications of reading disability subtypes so defined. By examining the subtypes within the context of a reading-level match, by looking for convergence with the similar investigation of Manis et al. (1996), and by reanalyzing the Castles and Coltheart data, we have distilled a consistent picture of developmental deviancy and developmental lag that appears to characterize the phonological and surface subtypes.

Phonological dyslexia is the more reliably identified subtype (i.e., if multiple indicators are used, they converge on a set of core phonological cases), and it appears to reflect true developmental deviancy. In other words, the pattern of

Table 5
Mean Differences Between the Phonological Dyslexics (N = 17) and the Reading-Level Controls (N = 23)

Variable	Dyslexics	Controls	*t* value
WRAT–R Reading[a]	49.9	51.0	−0.58
WRAT–R Spelling[a]	32.0	31.4	0.50
Woodcock Word Identification[a]	46.8	42.6	1.28
Woodcock Word Attack[a]	7.9	15.5	−5.07***
Exception words	8.3	6.4	2.04*
Regular words	7.2	9.2	−1.68
Pseudowords	9.9	16.8	−5.71***
Rosner AAT	13.9	21.1	−3.16**
Wordlikeness choice	12.2	11.8	0.51
Syntactic processing[b]	−.473	.235	−3.15**
Working memory[b]	−.172	.310	−2.30*

Note. WRAT–R = Wide Range Achievement Test—Revised; Rosner AAT = Rosner Auditory Analysis Test.
[a] raw score. [b] *z* score.
$*p < .05$. $**p < .01$. $***p < .001$, all two-tailed.

Table 6
Mean Differences Between the Phonological Dyslexics (N = 17) and the "Deviant on Both" Group (N = 19)

Variable	Phonological	Both	*t* value
WRAT–R Reading[a]	49.9	47.1	1.69
WRAT–R Spelling[a]	32.0	30.7	1.18
Woodcock Word Identification[a]	46.8	40.6	2.43*
Woodcock Word Attack[a]	7.9	8.0	0.08
Exception words	8.3	3.9	6.23***
Regular words	7.2	4.4	2.85**
Pseudowords	9.9	7.7	1.58
Rosner AAT	13.9	13.3	0.32
Wordlikeness choice	12.2	10.7	1.86
Syntactic processing[b]	−.473	−.308	0.61
Working memory[b]	−.172	−.352	0.76

Note. WRAT–R = Wide Range Achievement Test—Revised; Rosner AAT = Rosner Auditory Analysis Test.
[a] raw score. [b] *z* score.
* $p < .05$. ** $p < .01$. *** $p < .001$, all two-tailed.

linguistic and information processing strengths and weaknesses displayed by this subtype do not match those found in reading-level controls. In contrast, surface dyslexia has consistently (see Manis et al., 1996, and the previous reanalysis of Castles & Coltheart) resembled a form of developmental delay. It is interesting that when trying to simulate surface dyslexia with a connectionist network, Plaut and Shallice (1994) reported that damaging the network did not work as well as simply examining the undamaged network at an earlier point in its learning curve. They found that "a much better match to fluent surface dyslexia is found in the behavior of the *undamaged* network earlier in learning, before it has mastered the entire training corpus" (p. 24).

In light of the similarity in the data patterns of the RL comparisons in all three of these studies, it is interesting to note that the original Bryant and Impey (1986) study—the first to analyze the patterns revealed in the surface and phonological dyslexic case studies in the context of reading-level controls—obtained a converging outcome. The one pattern of HM (the phonological dyslexic of Temple & Marshall, 1983) that Bryant and Impey could not match to a child in their RL control group was HM's nonword reading. HM read considerably fewer nonwords than the worst nonword reader in Bryant and Impey's control group. This contrasts with the performance of CD, Coltheart et al.'s (1983) surface dyslexic. Bryant and Impey were able to find a match in their RL control group for every pattern displayed by CD. The performance of CD disappeared into the distribution of RL controls, as did that of most of the surface dyslexics in our study. Likewise, the four acquired surface dyslexia cases studied by Birnboim (1995) displayed many performance similarities to second-grade reading-level controls. In short, the results from case studies of developmental phonological and surface dyslexia are entirely consistent with the patterns displayed in three studies with larger scale sampling of reading-disabled children. Finally, the picture of the two subtypes defined from the converging studies discussed here bears a strong resemblance to the subtypes derived from Frith's (1985) developmental model of reading.

Consider how the two subtypes might arise through different combinations of relative phonological impairment and experience with print. Individuals who are matched on their level of phonological skill may vary greatly in their level of print exposure (Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1990; Stanovich & West, 1989). Low print exposure might not have very dire consequences for a reader with high levels of phonological coding skill. When such a reader does open a book, phonological coding enables the reading process, irrespective of the inadequately developed orthographic lexicon. However, the situation is probably different for a reader with somewhat depressed phonological skills (and one must never forget that even the surface dyslexics have phonological processing problems to some degree). Without efficiently functioning phonological coding processes, a system designed for compensatory processing would actually draw more on orthographic knowledge; however, in the case of the surface dyslexic, that orthographic knowledge may be lacking, in part because of inadequate exposure to print. For example, our investigation revealed a greater proportion of surface dyslexics—compared with phonological dyslexics—than observed by Castles and Coltheart (1993). The particular schools from which we drew our sample—with their low level of achievement and diverse population—might have, in part, accounted for our relatively high proportion of surface dyslexics. If lack of exposure to print at home and in school is, in part, responsible for the surface pattern and if the low achievement in our schools was, in fact, a proxy for inadequate exposure in the home (and perhaps in the school itself), then a greater proportion of surface cases would be expected in our setting.

Thus, surface dyslexia may arise from a milder form of phonological deficit than that of the phonological dyslexic, but one conjoined with exceptionally inadequate reading experience. This is a somewhat different interpretation of surface dyslexia (see also Manis et al., 1996) than the common one of differential impairment in a dual-route architecture (Castles & Coltheart, 1993). As Snowling, Bryant, and Hulme (1995) noted, "Many poor readers have low levels of exposure to print (Stanovich, 1993a)—lack of reading experience may cause dyslexic children to resemble surface dyslexic patients. Arguably, what such children lack is the word-specific knowledge that is normally acquired by reading. In our view, it is misleading to describe such children as having an aberrant 'lexical' but intact 'sub-lexical' mechanism" (p. 6).

In contrast, the phonological dyslexic pattern might become more apparent when a more severe pathology underlying the functional architecture of phonological coding (Castles & Coltheart, 1993; Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, et al., 1980, 1983; Patterson et al., 1985; Plaut & Shallice, 1994) is conjoined with relatively high levels of exposure to print. The latter would hasten the development of the orthographic lexicon (which is critical for the processing of exception words), but the former would be relatively refractory to direct remediation efforts

(Lovett et al., 1994; Lovett, Warren-Chaplin, Ransby, & Borden, 1990; Vellutino et al., 1996) and result in relatively slow growth in the ability to read pseudowords (Manis et al., 1993; Olson, 1994; Snowling, 1980).

We would conjecture that the two subtypes might separate when other methods of differentiating subtypes are used, for example, response to treatment, genetic analyses, and neurological investigation. Put another way, it is hypothesized that phonological dyslexia will be more refractory to treatment than surface dyslexia (see Vellutino et al., 1996), will have a higher heritability, and will more clearly display brain anomalies. On the basis of our data, we also conjectured that the "deviant on both" subtype would be more similar to phonological dyslexia than to surface dyslexia in these characteristics. Recall that this subtype comprised almost a third of our sample and that its cognitive profile was quite similar to that of the phonological dyslexics, except that the latter read exception and regular words somewhat more accurately. We conjectured that the overrepresentation of the former subtype in our sample (compared with the other two investigations) was because of the younger age of our sample.

How might a younger child deviant on both stimulus types develop into a phonological dyslexic? Some children in the both-deviant group might continue to practice reading and to receive considerable exposure to print (Stanovich, 1993a; Stanovich & Cunningham, 1993; Stanovich, West, & Harrison, 1995). This print exposure may result in these children having relatively less seriously impaired orthographic processing mechanisms (Siegel et al., 1995; Stanovich & Siegel, 1994; Zivian & Samuels, 1986). It may also result in these children building exception word recognition abilities (which depend on orthographic representations in the mental lexicon; see Ehri, 1992; Perfetti, 1992, 1994; Stanovich, 1990; Stanovich & West, 1989). However, their more seriously impaired phonological processing abilities will probably not develop at the same rate (Manis, Custodio, & Szeszulski, 1993; Olson, 1994; Snowling, 1980), thus resulting in greater dissociation between phonological coding ability and exception word fluency with development (see also Manis et al., 1996).

Only a longitudinal investigation can address this conjecture about the development of the children deviant on both stimulus types. Such investigations would also provide evidence relevant to the evaluation of the two major subtypes. In a similar manner, it might prove worthwhile to incorporate the surface–phonological subtype dimension into studies of treatment, genetic etiology, and brain structure–function correlates.

A Final Caveat on the Castles and Coltheart (1993) Methodology

One final point that needs to be emphasized is that the patterns and trends in the samples of reading-disabled children that have been considered here are, in some sense, a necessary consequence of the well-known finding that dyslexics show a pseudoword reading deficit in an RL match and simultaneously show no deficit in reading exception words (or in other tasks that rely heavily on orthographic coding; see Olson et al., 1989; Siegel et al., 1995; Stanovich & Siegel, 1994; Zivian & Samuels, 1986). When these empirical trends are put together with the innovative Castles and Coltheart (1993) procedure for defining subtypes, then it is almost necessarily the case that the trends that we have outlined will be found. Indeed, the conclusions about these different subtypes have actually been implicit in the findings of many studies using an RL match and examining pseudoword and exception word processing but have remained implicit until the Castles and Coltheart procedures revealed a way of drawing this implicit pattern to our attention. We demonstrated this point by running simulations based on data extrapolated from our RL controls.

For example, in one simulation, the data from the 44 CA controls and 23 RL controls in our sample were used without change. However, the data from the 68 dyslexics were simulated on the basis of data from the RL controls. Dyslexic pseudo-participants were randomly assigned exception word values from a distribution with the same mean and variance as the RL controls. They were simultaneously assigned pseudoword values from a distribution with the same variance as the RL controls but with a mean 3 points lower (simulating the RL deficit on pseudowords). The correlation between pseudowords and exception words in the simulated data was slightly lower (.40) than that in the actual data (.53), but this does not materially affect the results.

By using the regression lines from the CA controls (i.e., the Castles–Coltheart procedure), we identified substantial numbers of phonological and surface dyslexics in the simulated data set. When the RL controls are used to construct the regression lines, just as in the three actual empirical studies discussed here, 11 phonological dyslexics are still identified. In contrast, just as in the actual empirical data, surface dyslexia disappears when confidence intervals are derived from an RL rather than CA control group. (No surface dyslexics were identified in the simulated data.) This, then, is the sense in which we mean that these subtype patterns have been implicit in the statistical relationships among word recognition variables that have been known for some time. Virtually any data set that displays this pattern of relationships will reveal subtypes of poor readers with the characteristics that have been described by Castles and Coltheart (1993), Manis et al. (1996), and our investigation.

References

Adams, M. J., & Bruck, M. (1993). Word recognition: The interface of educational policies and scientific research. *Reading and Writing: An Interdisciplinary Journal, 5,* 113–139.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386–393.

Beauvois, M. F., & Derouesne, J. (1979). Phonological alexia: Three dissociations. *Journal of Neurology, Neurosurgery, and Psychiatry, 42,* 1115–1124.

Beech, J., & Harding, L. (1984). Phonemic processing and the poor reader from a developmental lag viewpoint. *Reading Research Quarterly, 19,* 357–366.

Berndt, R., Reggia, J., & Mitchum, C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers, 19,* 1–9.

Besner, D., Twilley, L., McCann, R., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432–446.

Birnboim, S. (1995). Acquired surface dyslexia: The evidence from Hebrew. *Applied Psycholinguistics, 16,* 83–102.

Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design study of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development, 63,* 999–1011.

Bradley, L., & Bryant, P. E. (1978). Difficulties in auditory organization as a possible cause of reading backwardness. *Nature, 271,* 746–747.

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

Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Child Psychology, 50,* 156–178.

Bryant, P. E., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen, and Ferguson. *Psychological Bulletin, 100,* 101–103.

Bryant, P., & Impey, L. (1986). The similarities between normal readers and developmental and acquired dyslexics. *Cognition, 24,* 121–137.

Byrne, B., Freebody, P., & Gates, A. (1992). Longitudinal data on the relations of word-reading strategies to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 27,* 141–151.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1–82). Orlando, FL: Academic Press.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47,* 149–180.

Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54,* 74–89.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589–608.

Coltheart, V., & Leahy, J. (1992). Children's and adult's reading of nonwords: Effects of regularity and consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 718–729.

Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *Quarterly Journal of Experimental Psychology, 35A,* 469–496.

Coltheart, M., Patterson, K., & Marshall, J. C. (1980). *Deep dyslexia.* London: Routledge & Kegan Paul.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82,* 733–740.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.

Ellis, A. W. (1979). Developmental and acquired dyslexia: Some observations on Jorm. *Cognition, 7,* 413–420.

Ellis, A. W. (1984). The cognitive neuropsychology of developmental (and acquired) dyslexia: A critical survey. *Cognitive Neuropsychology, 2,* 169–205.

Fletcher, J. M. (1992). The validity of distinguishing children with language and learning disabilities according to discrepancies with IQ: Introduction to the Special Series. *Journal of Learning Disabilities, 25,* 546–548.

Fletcher, J. M., & Morris, R. (1986). Classification of disabled learners: Beyond exclusionary definitions. In S. J. Ceci (Eds.), *Handbook of social and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 55–80). Hillsdale, NJ: Erlbaum.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D., Katz, L., Liberman, I., Stuebing, K., Francis, D. J., Fowler, A., & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86,* 6–23.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28,* 3–71.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301-330). London: Erlbaum.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove, England: Erlbaum.

Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology, 63,* 563–582.

Hanley, J. R., Hastie, K., & Kay, J. (1992). Developmental surface dyslexia and dysgraphia: An orthographic processing impairment. *Quarterly Journal of Experimental Psychology, 44A,* 285–319.

Hatcher, P., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65,* 41–57.

Hinton, G., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behavioral and Brain Sciences, 8,* 689–740.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery Program. *Journal of Educational Psychology, 85,* 112–126.

Jastak, S., & Wilkinson, G. S. (1984). *The wide range achievement test—Revised.* Wilmington, DE: Jastak Associates.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243–255.

Laxon, V., Smith, B., & Masterson, J. (1995). Children's nonword reading: Pseudohomophones, neighborhood size, and priming effects. *Reading Research Quarterly, 30,* 126–144.

Lovett, M. W., Borden, S., DeLuca, T., Lacerenza, L., Benson, N., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. *Developmental Psychology, 30,* 805–822.

Lovett, M., Warren-Chaplin, P., Ransby, M., & Borden, S. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects. *Journal of Educational Psychology, 82,* 769–780.

Lyon, G. R. (1987). Learning disabilities research: False starts and broken promises. In S. Vaughn & C. S. Bos (Eds.), *Research in learning disabilities* (pp. 69–85). Boston, MA: College-Hill Press.

Manis, F. R., Custodio, R., & Szeszulski, P. A. (1993). Development of phonological and orthographic skill: A 2-year longitudinal study of dyslexic children. *Journal of Experimental Child Psychology, 56,* 64–86.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Peterson, A. (1996). On the bases of two subtypes of developmental dyslexia. *Cognition, 58,* 157–195.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2,* 175–199.

McBride-Chang, C. (1995a). Phonological processing, speech perception, and reading disability: An integrative review. *Educational Psychologist, 30,* 109–121.

McBride-Chang, C. (1995b). What is phonological awareness? *Journal of Educational Psychology, 87,* 179–192.

McCloskey, M. (1991). Networks and theories: The place of connectionism in cognitive science. *Psychological Science, 2,* 387–395.

McKinney, J. (1984). The search for subtypes of specific learning disability. *Journal of Learning Disabilities, 17,* 43–50.

Metsala, J. L., & Brown, G. D. A. (in press). Normal and dyslexic reading development: The role of formal models. In C. K. Leong & C. Hulme (Eds.), *Cognitive and linguistic bases of reading and spelling.* Hillsdale, NJ: Erlbaum.

Morris, R., & Satz, P. (1984). Classification issues in subtype research: An application of some methods and concepts. In R. Malatesha & H. Whitaker (Eds.), *Dyslexia: A global issue* (pp. 59–82). The Hague: Martinus Nijhoff.

Murphy, L., & Pollatsek, A. (1994). Developmental dyslexia: Heterogeneity without discrete subgroups. *Annals of Dyslexia, 44,* 120–146.

Olson, R. K. (1994). Language deficits in "specific" reading disability. In M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 895–916). San Diego, CA: Academic Press.

Olson, R. K., Forsberg, H., & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V. Berninger (Eds.), *Varieties of orthographic knowledge: Theoretical and developmental issues* (Vol. 1, pp. 27–71). Dordrecht, The Netherlands: Kluwer Academic.

Olson, R. K., Wise, B., Conners, F., & Rack, J. (1990). Organization, heritability, and remediation of component word recognition and language skills in disabled readers. In T. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 261–322). New York: Academic Press.

Olson, R. K., Wise, B., Conners, F., Rack, J., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities, 22,* 339–348.

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12, pp. 421–447). London: Erlbaum.

Patterson, K., Marshall, J. C., & Coltheart, M. (1985). *Surface dyslexia.* London: Erlbaum.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 335–359). London: Erlbaum.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 849–894). San Diego, CA: Academic Press.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Plaut, D., & Shallice, T. (1994). *Connectionist modelling in cognitive neuropsychology: A case study.* Hove, England: Erlbaum.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27,* 28–53.

Rosner, J., & Simon, D. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities, 4,* 40–48.

Satz, P., Morris, R., & Fletcher, J. M. (1985). Hypotheses, subtypes, and individual differences in dyslexia: Some reflections. In D. B. Gray & J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 25–40). Parkton, MD: New York Press.

Seidenberg, M. S. (1993). A connectionist modeling approach to word recognition and dyslexia. *Psychological Science, 4,* 299–304.

Seidenberg, M. S. (1994). Language and connectionism: The developing interface. *Cognition, 50,* 385–401.

Seidenberg, M. S., & McClelland, J. L. (1989). Visual word recognition and pronunciation: A computational model of acquisition, skilled performance, and dyslexia. In A. M. Galaburda (Eds.), *From reading to neurons* (pp. 256–305). Cambridge, MA: MIT Press.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shankweiler, D., Crain, S., Brady, S., & Macaruso, P. (1992). Identifying the causes of reading disability. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 275–305). Hillsdale, NJ: Erlbaum.

Shankweiler, D., Crain, S., Katz, L., Fowler, A., Liberman, A., Brady, S., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J., Stuebing, K., Shaywitz, S., & Shaywitz, B. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science, 6,* 149–156.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education: Contributions from Educational Psychology, 1,* 1–57.

Shaywitz, B. A., Fletcher, J. M., Holahan, J. M., & Shaywitz, S. E. (1992). Discrepancy compared to low achievement definitions of reading disability: Results from the Connecticut Longitudinal Study. *Journal of Learning Disabilities, 25,* 639–648.

Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42,* 201–215.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22,* 469–479.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25,* 618–629.

Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology, 24,* 28–37.

Siegel, L. S., Share, D., & Geva, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6,* 250–254.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences, 11,* 1–74.

Smolensky, P. (1989). Connectionist modeling: Neural computation/mental connections. In L. Nadel, L. A. Cooper, P. Culicover, & P. M. Harnish (Eds.), *Neural connections, mental computation* (pp. 49–67). Cambridge, MA: MIT Press.

Snowling, M. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. *Journal of Experimental Child Psychology, 29,* 294–305.

Snowling, M. (1983). The comparison of acquired and developmental disorders of reading—A discussion. *Cognition, 14,* 105–118.

Snowling, M. J., Bryant, P. E., & Hulme, C. (1995). Theoretical and methodological pitfalls in making comparisons between developmental and acquired dyslexia: Some comments on Castles and Coltheart (1993). Unpublished manuscript.

Speece, D., & Cooper, D. (1991). Retreat, regroup, or advance? An agenda for empirical classification research in learning disabilities. In L. Feagans, E. Short, & L. Meltzer (Eds.), *Subtypes of learning disabilities* (pp. 33–49). Hillsdale, NJ: Erlbaum.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21,* 590–612.

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Developmental Review, 10,* 72–100.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26,* 7–29.

Stanovich, K. E. (1993a). Does reading make you smarter? Literacy and the development of verbal intelligence. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 133–180). San Diego, CA: Academic Press.

Stanovich, K. E. (1993b). Dysrationalia: A new specific learning disability. *Journal of Learning Disabilities, 26,* 501–515.

Stanovich, K. E. (1994). Does dyslexia exist? *Journal of Child Psychology and Psychiatry, 35,* 579–595.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology, 85,* 211–229.

Stanovich, K. E., Nathan, R. G., & Zolman, J. E. (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. *Child Development, 59,* 71–86.

Stanovich, K. E., & Siegel, L. S. (1994). The phenotypic performance profile of reading-disabled children: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402–433.

Stanovich, K. E., West, R. F., & Harrison, M. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology, 31,* 811–826.

Temple, C. M., & Marshall, J. C. (1983). A case study of developmental phonological dyslexia. *British Journal of Psychology, 74,* 517–533.

Torgesen, J. K. (1991). Subtypes as prototypes: Extended studies of rationally defined extreme groups. In L. Feagans, E. Short, & L. Meltzer (Eds.), *Subtypes of learning disabilities* (pp. 229–246). Hillsdale, NJ: Erlbaum.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology, 83,* 437–443.

Vellutino, F. R., Scanlon, D. M., Sipay, E., Small, S., Pratt, A., Chen, R., & Denckla, M. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Toward distinguishing between constitutionally and experientially based causes of reading disability. *Journal of Educational Psychology, 88,* 601–638.

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

Waters, G., & Seidenberg, M. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition, 13,* 557–572.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests—Revised.* Circle Pines, MN: American Guidance Service.

Zivian, M. T., & Samuels, M. T. (1986). Performance on a word-likeness task by normal readers and reading-disabled children. *Reading Research Quarterly, 21,* 150–160.

Received January 11, 1996
Revision received June 21, 1996
Accepted June 21, 1996 ■