Further Analysis of the Relationship Between Reading Achievement and Intelligence: Response to Naglieri

Frank R. Vellutino

Naglieri's response to the IQ-achievement discrepancy article written by myself and my colleagues (Vellutino, Scanlon, & Lyon, 2000) reflects a view regarding the relationship between intelligence and reading achievement long held by many scholars and practitioners—specifically, that the two are sufficiently correlated to justify the use of intelligence test scores to predict reading achievement and, thereby, to define reading disability and other learning disabilities (see Vellutino et al., 2000, for a review). It also reflects what appears to be anxious concern over the conclusions that might be drawn from the research reviewed and the results reported in our article. I have several reactions to the points made in Naglieri's commentary.

Naglieri seems to perceive as the main message of our article the assertion that reading achievement in the general sense of the term is not reliably or strongly correlated with intelligence. If this were the main message of our article (Vellutino et al., 2000), then his arguments would have greater force. However, in no place in our article did we make such a blanket assertion. Instead, we asserted that when reading achievement is defined in terms of measures of basic reading subskills such as word identification and letter-sound decoding, there is reasonably strong evidence to suggest that correlations between tests of intelligence and tests evaluating such skills tend to be mercurial and modest in size at best. This, of course, suggests that the logic behind the use of the IQ-achievement discrepancy to define reading disability is flawed because discrepancy scores are typically derived from tests of basic reading skills, and the evidence suggests that a child does not have to have a very high IQ to score at least in the average range on such tests. Indeed, we cited a considerable amount of evidence to support this assertion, including results from a study by Siegel (1988) that showed that even children with marginal IQs (75-85) were able to score in at least the average range on tests of word identification and pseudoword decoding. Similar results were obtained in a study by Share, McGee, and Silva (1989), which we also discussed. Moreover, in an earlier study (Vellutino et al., 1996) that was the primary focus of our recent article (Vellutino et al., 2000), poor readers at the beginning stages of reading development were found to be equivalent to average-IQ normal readers on measures of both verbal and nonverbal intelligence, and neither of these measures distinguished between tutored children who were found to be difficult to remediate and tutored children who were found to be readily remediated. At the same time, average- and above-average-IQ normal readers were found to be equivalent on measures of word identification and letter-sound decoding. Thus, there is strong evidence for the assertion that measures of intelligence are only weakly to moderately correlated with measures of basic reading subskills.

However, we (Vellutino et al., 2000) were careful to point out that when reading achievement is defined as the ability to comprehend written text, correlations between measures of intelligence and measures of reading achievement tend to be more robust and less mercurial, because the various intelligence tests available typically contain items that evaluate knowledge and abilities that are entailed on measures of reading comprehension—for example, vocabulary knowledge, analytic skills, attention to detail, inferencing, reasoning, and so forth. The point here is that Naglieri's commentary does not make a clear enough distinction between the two definitions of reading achievement discussed in the Vellutino et al. (2000) article. This is important because most of the research cited in his commentary, supporting what he believes to be a strong relationship between reading achievement and measured intelligence, is based on correlations between measures of reading comprehension and measures of intelligence. This is certainly true of the research he has conducted, and in the article by Stanovich, Cunningham, and Feeman (1984), the weak to moderately high correlations between measures of reading achievement and measures of intelligence reported by these investigators (Stanovich et al., 1984, Table 1) are by and large based on group-administered tests of reading comprehension and group-administered tests of intelligence. Thus, most of the arguments presented by Naglieri simply do
not apply to our article (Vellutino et al., 2000).

A related point I wish to make is concerned with the evidence Naglieri presented for his contention that intelligence and reading achievement are reliably and strongly correlated. He first presented evidence based on correlations involving group-administered reading and intelligence tests, but as we pointed out (Vellutino et al., 2000) and he acknowledged, most group-administered intelligence tests require reading ability and, therefore, confound intelligence and reading ability. Yet despite the fact that most of the reading tests employed in the studies discussed in Naglieri's commentary evaluated reading comprehension, the lion's share of the observed correlations were modest in size (average or median correlations, ranging from .50 to .57), including those from studies using nonverbal tests of intelligence such as the Naglieri Nonverbal Ability Test (NNAT; Naglieri, 1997), which, Naglieri suggested, avoid the problems associated with group-administered tests of intelligence that require reading ability.

Naglieri then presented evidence based on correlations between individually administered tests of reading achievement and composite scores from the Cognitive Assessment System (CAS; Das & Naglieri, 1997) and the K-ABC battery (Kaufman & Kaufman, 1983), which he claimed are ideal for measuring intellectual abilities because they do not include “subtests (e.g., Vocabulary) that are obviously related to reading achievement.” As regards the Cognitive Assessment System, Naglieri reported correlations between the CAS total test standard score and the following measures of basic reading subskills from the Woodcock-Johnson battery (Woodcock & Johnson, 1989): Broad Reading ($r = .71$); Reading Comprehension ($r = .72$); Letter-Word Identification ($r = .66$); Word Attack ($r = .62$); and Basic Reading Skills (a composite of the latter two, $r = .69$). However, all of these correlations are based on a sample ($n = 1,600$) of children between the ages of 5 and 17 and, therefore, mask age/grade differences in the relationship between the CAS and the various measures of reading achievement. This is problematic because, as documented by Stanovich et al. (1984), correlations between measures of reading achievement and measures of intelligence tend to be lower at the early stages of reading development than at later stages of reading development. Stanovich et al. (1984) also suggested that this phenomenon is a consequence of “reciprocal causation” between reading development and intellectual development, which results in what Stanovich has called “Matthew effects” (Stanovich, 1986, p. 360), that is, the deleterious effects of long-standing reading disorder on intelligence test measures that depend in part on reading ability (e.g., measures evaluating vocabulary knowledge or general information). The result is that correlations between reading achievement and measures of intelligence are often artificially inflated in older children, and it is possible that such effects obtain in the case of correlations between scores derived from tests of basic reading skills and scores derived from tests included in the CAS battery.

To make this argument more concrete, a brief perusal of Table 4.17 in the CAS interpretive manual (Das & Naglieri, 1997, p. 63) shows that correlations between subtests of the CAS battery (Planning, Attention, Simultaneous, Successive) and the Woodcock-Johnson Letter-Word Identification and Word Attack subtests are generally lower for children between the ages of 5 and 10 compared with those for children between the ages of 11 and 17. (Sample sizes are $n = 630$ and $n = 454$ for children in the 5-7 and 8-10 age ranges, respectively, and $n = 228$ for children in both the 11-13 and 14-17 age ranges). Moreover, these correlations range from .31 to .66, and 6 out of 8 of them are below .50. At the same time, correlations between the CAS measures and the Woodcock-Johnson Passage Comprehension subtest range from .31 to .66, and 6 out of 8 of them are below .50.

These results are quite in keeping with the research we reviewed (Vellutino et al., 2000) documenting that correlations between measures of reading achievement and measures of intelligence were found to be quite modest in samples of early elementary school-age children and ranged only from .40 to .60. Thus, in accord with results obtained in previous research, CAS sub-scales account for no more than 36% of the variance on measures of basic reading subskills administered to early elementary-age children and, in most cases, account for considerably less. Even if we include the data from children in the older age groups (children between ages 11 and 17), the range for these correlations is only between .31 and .66, and correlational averages across all these ages and all test measures range only from .39 to .54.

Finally, all estimates presented in Table 4.17 of the CAS interpretive manual (Das & Naglieri, 1997, p. 63) are based on a range of ages and may themselves be artificially inflated especially at the lower age levels. Indeed, because of formal exposure to reading instruction and greater experience in reading, the 7-year-old child is apt to have acquired greater proficiency in word identification and word decoding than the 5-year-old child, and the 10-year-old child is apt to have acquired greater proficiency in these skills than the 8-year-old child; this suggests that variability produced by the reading measures would be greater among children across the age levels included in a given age range than among children at one or another of the age levels within that range. Thus, among children at a given age level (e.g., 5-year-olds), the correlations between the CAS subtests and measures of basic reading skills may even be smaller than those presented in Table 4.17 for children in the age range that included that age level (e.g., 5- to 7-year-olds).

We should also point out that the correlations between the CAS subtests and measures of reading achievement
are even lower for individuals with learning disabilities (n = 81). Thus, for children from this population, the range of correlations between the CAS subtests and Basic Reading Skills is between .27 and .38, whereas the range for Reading Comprehension is between .33 and .45 (see Das & Naglieri, 1997, Table 4.19, p. 68). This is the same pattern of results that emerged in our own study with comparable size samples (see Vellutino et al., 2000, Table 2, p. 228; Table 4, p. 232).

As regards the K-ABC battery, Naglieri has pointed out that the mental processing composite on this battery correlated .63 with the Passage Comprehension subtest of the Woodcock-Johnson battery for children between the ages of 5 and 12, which by his standards is quite large. What he failed to point out is that the average (raw score) correlations between the K-ABC Mental Processing subtests and tests of Reading/Decoding for children between the ages of 5 and 12 inclusive (n = 1,500) range from .30 to .53, and the majority of these correlations are below .45. At the same time, the range of correlations between the same K-ABC subtests and tests of Reading/Understanding is from .27 to .46, and the majority fall below .40. When we examine correlations between these same measures for the age groups separately (sample sizes ranging from 100 to 200), we see that they tend to be lower in children close to the beginning stages of reading than in children in the older age groups, consistent with the pattern observed for the CAS subtests. Moreover, the majority of the correlations within each age group are no better than .45 in the case of the Reading/Decoding tests, which means that in the majority of instances, the mental ability tests included in the K-ABC battery account for no more than 20% of the variance on tests of basic reading subskills. Thus, results from the K-ABC battery, like results from the CAS battery, provide a weak foundation for the claim that basic reading subskills such as word identification and word decoding are strongly and reliably correlated with tests of intelligence.

A further point I wish to make relates to Naglieri’s claim that intelligence and reading achievement are strongly and reliably correlated as evidenced by results of nonverbal intelligence tests such as the Naglieri Nonverbal Ability Test (Naglieri, 1997) and individually administered tests such as the CAS and the K-ABC, which he has suggested are ideal for evaluating intellectual abilities in developing readers because they do not include subtests that entail knowledge and abilities that are related to reading achievement. We have already documented that all three of these measures produced weak to modest correlations even when the reading achievement measure involved reading comprehension; this by itself vitiates Naglieri’s claim. However, it is also important to point out that, whereas nonverbal ability tests such as the NNAT do by and large reduce confounding by reading-related cognitive abilities, the same cannot be said of all of the CAS and the K-ABC subtests. Both the CAS and the K-ABC batteries include subtests that evaluate memory for the order of verbally encoded stimuli (e.g., word order, order of digits, sentence repetition), and each of these tasks places heavy demands on verbal short-term memory. This is noteworthy because verbal memory tasks have been found to be significantly and reliably correlated with reading achievement and produce statistically significant differences between poor and normal readers (Katz, Shankweiler, & Liberman, 1981; see also Vellutino et al., 1996, for a review of relevant research). Similarly, the Planning subtest of the CAS requires familiarity with letters of the alphabet and familiarity with digits from 1 to 9, and it has been shown that such knowledge reliably predicts success in beginning reading (Vellutino et al., 1996). This subtest also produces statistically significant differences between poor and normal readers. Moreover, the Attention subtest of the CAS actually requires reading of color words, and even the so-called “nonverbal” subtests of both the CAS and K-ABC batteries—that is, those designed to evaluate what has come to be called simultaneous processing—are confounded by verbal processing ability. For example, the Verbal-Spatial processing subtest of the Simultaneous Processing component of the CAS requires that the examinee carefully listen to and carry out oral directions, which is a skill that has been shown to reliably distinguish between poor and normal readers (Vellutino et al., 1996). Similarly, the Gestalt Closure subtest of the Simultaneous Processing component of the K-ABC requires that the child name objects or scenes pictured in fragmented form, and there is now abundant evidence that name encoding and retrieval are cognitive abilities that are significantly correlated with reading achievement and reliably distinguish between poor and normal readers (see Vellutino et al., 1996, for a review). Thus, Naglieri’s claim that the CAS and the K-ABC are ideal for evaluating the intellectual abilities of developing readers because they do not include subtests that are related to reading achievement is arguable. Indeed, because subtests on both the CAS and the K-ABC batteries actually do entail reading-related cognitive abilities, correlations between measures of reading achievement and measures of performance on either of these batteries may not be trustworthy and should be interpreted with the utmost caution, especially if they are used to define reading disability. The use of composite scores from either of these batteries is especially problematic because composite scores mask the contribution made by reading-related cognitive abilities to correlations involving measures of reading achievement.

The preceding discussion underscores a more general problem with the use of intelligence tests to predict reading achievement, which is that the types of skills evaluated by and included on the plethora of tests that might be used for this purpose are not unpacked. We know something about why some types of intelligence test items correlate reliably with measures of reading achievement, but not most of them. For example, many intelli-
gence tests include measures of vocabulary knowledge, and there is considerable evidence that such measures correlate very highly with measures of language comprehension and, thereby, with measures of reading comprehension. This is no doubt because someone with a reasonably large vocabulary has a better chance of comprehending more of what he or she reads than someone with a small vocabulary. Analogously, what can Naglieri say about why scores on tests such as the NNAT, CAS, and K-ABC should be expected to correlate with basic reading skills, such as word identification and letter-sound decoding, or reading-related skills, such as phonemic awareness, that are known to be causally related to the acquisition of such basic reading skills? It is not enough to assert that such measures should be expected to correlate highly with reading and reading-related skills because they measure intelligence simply on the strength of the traditional assumption that highly intelligent people should learn to read better than less intelligent people. If items such as those included on the NNAT, the CAS, and the K-ABC reliably and strongly correlate with measures of reading achievement, including measures of basic subskills such as word identification or word decoding (and I do not believe that this is the case, as is evident from the results cited previously), the important theoretical question to ask is why this should be so. To simply state that the measures in question are reliably and strongly correlated with reading achievement because they evaluate intelligence is theoretically opaque and circular, especially when the question of exactly what these tests are measuring is an open one, as I have tried to point out (see exchanges between Kranzler & Keith, 1999a, 1999b, and Naglieri, 1999, for more on this question).

A final point I wish to make has to do with Naglieri’s argument that the correlations cited in both Stanovich et al.’s (1984) and Vellutino et al.’s (2000) articles were based on small samples and are therefore flawed. As pointed out by Stanovich et al. (1984), the limitations imposed by small samples are often offset by the reliability of given findings, and this was the case in both studies. It can be seen in Table 1 of the Vellutino et al. article that the magnitudes of the correlations between the intelligence and the reading measures with sample sizes of 247 and 171 participants for the two respective age/grade groups evaluated were comparable to the magnitudes of the correlations obtained between similar measures with much smaller sample sizes, which speaks for the reliability of the correlations. Moreover, these correlations tend to be weak, especially those involving the test of nonverbal intelligence (Wechsler Intelligence Scale for Children-Revised, Performance subscale; Wechsler, 1974). In contrast, Naglieri has reported one set of correlations involving the CAS and measures of basic reading skills, and he suggested that because the sample is based on a little more than 1,600 participants, his results are more reliable than those reported by Stanovich et al. and Vellutino et al. and, therefore, make a compelling case for his claim that reading achievement and intelligence are strongly and reliably correlated.

This assertion can be challenged. First, as I indicated earlier, the correlations Naglieri reported are confounded by age differences, they are modest in size, and they are substantially lower for early elementary–age children than for middle school– and high school–age children. Second, the sample size figures cited by Naglieri (N = 1,600) is misleading because there are at least three age groups within each of the age ranges presented in Table 4.17 (Das & Naglieri, 1997, p. 63). Although the total number of participants in each of these age groups is not presented, the total in each group can only be a fraction of the totals presented for each of the age ranges in the table (see previous discussion for the totals in each age range); it is therefore unlikely that any of these samples are substantially larger than either of the samples on which the data presented in Table 1 in Vellutino et al.’s (2000) article are based.

Finally, even if one were to find that the average correlation between tests such as the CAS, the K-ABC, and the NNAT on the one hand and tests of basic reading subskills on the other hand was something like .60 (which is unlikely, given the range of correlations between such measures reported in this and in Naglieri’s commentary), this would mean that tests of the former type would on average account for about one third of the variance on tests of the latter type, which is hardly a basis for claiming that tests of intelligence should be used to determine whether a child is working up to his or her potential in acquiring beginning literacy skills. This, of course, is the central point made in our original article (Vellutino et al., 2000) and is actually at the heart of current debate surrounding the use of the IQ–achievement discrepancy definition of specific reading disability.

In conclusion, I believe the arguments presented by Naglieri are oblique to the central theme of our original article (Vellutino et al., 2000), and most of the data he cites supporting what he believes to be a strong and reliable relationship between reading achievement and measured intelligence do not apply to the article because they involve correlations between measures of reading comprehension and measures of intelligence. Moreover, he has not made a convincing case for his claim that measures of intelligence on the one hand and measures of basic reading subskills such as word identification and letter-sound decoding on the other hand are strongly and reliably correlated. In fact, there is abundant evidence to suggest otherwise, including data published in the CAS interpretive manual (Das & Naglieri, 1997) and Kaufman’s K-ABC technical manual (Kaufman & Kaufman, 1983), which in both instances reflect a range of weak to moderate correlations that is quite in keeping with the range of weak to moderate correlations cited in our own article (Vellutino et al., 2000). Thus, there are sufficient grounds for questioning the use of the IQ–achievement discrep-
ancy to define specific reading disability.

ABOUT THE AUTHOR

Frank R. Vellutino, PhD, currently holds joint faculty appointments in the Department of Psychology (Cognitive Psychology Program), the Department of Educational and Counseling Psychology, and the Program in Linguistics and Cognitive Science of the Department of Anthropology at The University at Albany, State University of New York. His research is concentrated in related areas such as cognitive psychology, psycholinguistics, and developmental psychology, and he is especially interested in the cognitive and linguistic underpinnings of reading development and reading disability. Address: Frank R. Vellutino, Child Research and Study Center, Husted 134, The University at Albany, 135 Western Avenue, Albany, NY 12222.

REFERENCES