More Smoke and Mirrors: A Critique of the National Reading Panel Report on Fluency

At worst, the impact of free reading appears to be the same as that of traditional instruction, and it is often better, especially when studies are continued for more than an academic year, a finding that the National Reading Panel has obscured by omitting important studies and by describing others incorrectly, Mr. Krashen charges.

BY STEPHEN KRASHEN

In her review of the report on phonics by the National Reading Panel (NRP), Elaine Garan concluded that the report involved "a limited number of studies of a narrow population." I will argue that this problem is not limited to the section on phonics; it also applies to the section on the report on fluency. It is only by omitting a large number of relevant studies — and misinterpreting the ones that were included — that the NRP was able to reach the startling conclusion that there is no clear evidence that encouraging children to read more actually improves reading achievement.

OMISSIONS

The selection criteria used by the NRP to choose studies for its report were as follows:

1. The study had to be a research study that appeared to consider the efect of encouraging students to read more on reading achievement.
2. The study had to focus on English reading education, conducted with children (K-12).
3. The study itself had to have appeared in a refereed journal.
4. The study had to have been carried out with English language reading.

The NRP claimed that it could find only 14 studies that met these criteria. Of these, 10 were studies of the impact of sustained silent reading (SSR) programs in which some class time is set aside for free reading with little or no "accountability," and the missing 4 had positive results, with the students who were engaged in free voluntary reading outperforming comparison groups. Another study showed positive results for one condition but not for other conditions, and the rest of the studies showed no difference between groups or no gains. Table 1 summarizes these outcomes.

In other sections of the NRP report, such as the sections on phonics and phonic awareness, the NRP listed studies that were excluded from its analysis. This was not done for the section on fluency. We do not know, therefore, which excluded studies were simply missed and which were rejected, nor do we know the specific rationale for their rejection.

In Table 2, I present an "expanded" set

---

**TABLE 1.**

Duration of Treatment and Outcomes of SSR Studies: NRP Set*

<table>
<thead>
<tr>
<th>Duration</th>
<th>Positive</th>
<th>No Difference</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than seven months</td>
<td>2</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Seven months to one year</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

*The table shows 14 comparisons derived from a total of 10 studies.
**Students in sustained silent reading programs outperform students in control groups.

---

STEPHEN KRASHEN is a professor of education at the University of Southern California, Los Angeles.
of SSR studies in which tests of reading comprehension were used. Many of the studies summarized in Table 2 meet the four criteria of the NRP and were apparently missed. But there were some "violations." A few were done with students slightly older than the age limit imposed by the NRP; in all such cases, the subjects were undergraduate college students. Subjects in some of the studies were students of English as a second language. In several studies, students read in Spanish, not English; in these cases, the students were native speakers of Spanish. Finally, some studies were not published in refereed journals.

Table 2 summarizes the results of these studies. It includes studies included by the NRP as well as those that the NRP did not include.*

In the studies in Table 2, SSR students did as well as or better than comparison students in 50 out of 53 comparisons. For longer-term studies (those longer than one year), SSR students were superior in eight of 10 studies, and there was no difference in the other two. Moreover, there are plausible reasons why the results were not even more positive. In one study by Isabel Schon, Kenneth Hopkins, and Carol Vojir, there was no difference between SSR students and comparison groups, but only five of the 11 SSR teachers actually carried out SSR conscientiously.* The classes taught by these five achieved significantly better gains.

In a study by Ruth Cline and George Kretke, another study showing no difference, the subjects were junior high school students who were reading two years above grade level and probably had already established a reading habit. Similarly, in Zephaniah Davis' study of eighth-graders, SSR helped medium-level readers but not better readers.** SSR appears to be most effective for less mature readers, its aim being to interest them in outside reading. Those who are already dedicated readers will not show dramatic gains. It is doubtful, for example, that readers of this article will show much improvement if they add to their daily schedule an extra 10 minutes of reading.

It is important to note that the NRP did not include any studies lasting longer than one year. A more comprehensive review of the literature indicates that the positive impact of recreational reading increases over time.

Even applying the NRP's stricter criteria, SSR does very well, with readers doing as well or better than controls in 35 out of 36 comparisons. This suggests that the "violations" do not affect the central issue of whether encouraging recreational reading affects literacy development. Even if one allows only those studies that strictly meet the NRP's criteria, the result still favors recreational reading.

** Misinterpreted Studies

In addition to excluding relevant studies, the NRP misinterpreted some of the studies that it did include. The study by Ronald Carver and Robert Liebert should not have been cited as evidence for or against recreational reading, because the students were constrained with respect to what they could read. They were allowed to read books only at or below their level, and the choice of books was limited to 135 titles (the regular library stacks were off limits). There was heavy use of extrinsic motivators, the students had to take multiple-choice tests on the books they read, and reading time was heavily concentrated, with students reading in two-hour blocks.** Successful SSR programs allow students to read any books they wish, do not use extrinsic motivators, do not make students accountable for what they read, provide a wide variety of books to choose from, and typically meet for a short time each day over a long period.***

The NRP claimed that the advantage shown by Joanne Burley's study** was "small." Students in SSR were clearly significantly better in reading comprehension than comparison students in three other conditions, but it was not possible to calculate measures of the size of the effect. It is not clear how the NRP concluded that the difference was small, especially considering the fact that the treatment lasted only six weeks and contained only 14 hours of reading. In a response to a commentary of mine, Shanahan claimed that "the problem here was not with the statistics, but with the design of the study. Each of the four treatments was offered by a different teacher, and students were not randomly assigned to the groups. It is impossible to unambiguously attribute the treatment differences to the methods."*** This is not accurate. Student assignment was in fact random, and each of the four teachers was randomly assigned to one of the four groups. In addition, the group that did SSR was superior to all three comparison groups, taught by three different teachers.

Not included in my summary of studies in Table 2 is a study that the NRP did include. Janet Langford and Elizabeth Allen used the Slossen Oral Reading Test, which consisted of reading words aloud, which may or may not involve genuine reading comprehension. The difference between the groups was highly significant, and I calculated an effect size of 1.055, which is quite large. Nevertheless, in discussing this study, the NRP concluded that "the gains were so small as to be of questionable educational value."***

In another study, students who engaged in SSR made better gains than a comparison group, but the difference was not statistically significant. SSR was significantly better than traditional instruction, however, when readers interacted with one another, that is, when they discussed their reading with one another and shared books. The NRP refers to this group's advantage.
as "slight," but it is not clear how the panel arrived at this conclusion. I computed a respectable effect size of .57 for the difference between the peer-interaction group and the comparison group.

The NRP interpreted a study by Sandra Holt and Frances O'Tuel as showing no difference between readers and comparison groups in reading comprehension. This study contained two samples: seventh- and eighth-graders. According to the text of the article, for the total sample, the readers were significantly better on tests of reading comprehension. The text also states that the difference was statistically significant for the seventh-graders but not the eighth-graders, a conclusion that is consistent with mean posttest scores presented in the researchers' Table 1 (posttest means were not presented). In their Table 2, however, the difference for reading comprehension for grade 7 was not statistically significant. The effect size for grade 7 (my calculations), based on posttest means, was a substantial .58. The NRP did not mention this discrepancy. I classified the results of this study as a split decision.

The NRP included a 1991 study and reported that the study found no differences between SSR and skills practice. What the panel did not mention is that the entire treatment lasted only 10 days (not one month, as the NRP reports) and that each of four skills groups did intensive work on specific comprehension skills (locating details, drawing conclusions, finding the main idea, finding the sequence). The researchers found no difference among the five groups on tests of comprehension skills and concluded that "engaging in sustained reading of connected and meaningful text appeared to be just as effective as spending time on the learning and practicing of discrete comprehension skills."  

**ADDITIONAL EVIDENCE**

I should also point out that the case for reading does not rest entirely on studies of sustained silent reading. In "read and test" studies, subjects show clear gains in vocabulary and spelling after a brief exposure to comprehensible text. It is hard to attribute these gains to anything but reading. In addition, there are compelling case histories that cannot be easily explained on the basis of a competing hypothesis—cases such as Richard Wright, who credits reading with providing him with high levels of literacy development: "I wanted to write, and I did not even know the English language. I bought English grammars and found them dull. I felt that I was getting a better sense of the language from novels than from grammars." Or consider Ben Carson, a neurosurgeon who says that his mother's insistence that he read two books (of his own choosing) each week when he was in the fifth grade was a turning point in his life. Carson credits reading with improving his reading comprehension, vocabulary, and spelling, as well as helping him move from the bottom of his class in grade 5 to the top in grade 7. "Yes, I know: there was no control group, no tests were given, and the results were not in a refereed journal. But it is hard to imagine any other source for this obvious improvement. And cases like these are not uncommon.

**CONCLUSIONS**

The NRP concluded that "the handful of experimental studies" in which encouraging voluntary reading has been done "raise serious questions" about its efficacy. There are more than a handful of studies. Moreover, the addition of more studies to the analysis provides substantial evidence in support of the effectiveness of recreational reading.

Note that even a finding of "no difference" between free readers and students in traditional programs suggests that free reading is just as good as traditional instruction, which confirms that free reading does indeed result in literacy growth, an important theoretical and practical point. Because free reading is so much more pleasant than regular instruction (for both students and teachers) and because it provides students with valuable information and insights, a finding of no difference provides strong evidence in favor of free reading in classrooms.
At worst, the impact of free reading appears to be the same as that of traditional instruction, and it is often better, especially when studies are continued for more than an academic year, a finding that the National Reading Panel has obscured by omitting important studies and by describing others incorrectly. Garam ask us to look beneath the smoke and behind the mirrors of the NRP report on phonics. The same needs to be done with the NRP report on ‘encouraging fluency.’

2. The report on fluency is one of the Reports of the Subgroups that make up the Report of the National Reading Panel: Teaching Children to Read. The document is available at www.nichd.nih.gov/publica-
tions/nrp/report.htm.
3. Ibid., p. 3-23.
4. The NRP report devoted only six pages to recreational reading. In contrast, 66 pages are devoted to phonemic awareness and nearly as many to phonics.
5. The studies included in Table 2 follow. Duration less than seven months, findings positive: Joan Allaby and John M. King, “Effects of Uninterrupted Sustained Silent Reading and of Reading Skills Instruction on Changes in Second-Year High School Students’ Reading Attitudes and Achievement,” in P. David Pearson and Jane Hanson, eds., 27th Yearbook of the National Reading Conference (Clemson, S.C.: National Reading Conference, 1978), pp. 226-8; Mabel Aran-
"I’m scared.

For some kids, school bullying is a harsh reality.
You can decrease aggressive behavior and increase social competency with...

PATHS

(Promoting Alternative Thinking Strategies)

A proven-effective, research-based violence prevention curriculum for grades K-6!

By teaching kids self-control and problem-solving skills, you can make sure your school will
see better behavior in the classrooms, the halls, the cafeteria, and the playground. Award-winning and easy to use,
PATHS means safer students.

FREE 30-day PATHS preview—
a $640 value.*

Call 1-877-896-8532 today!

Channing L. Bete Co., Inc.
www.channing-bete.com

*If you decide not to keep PATHS, we request that you pay return shipping.

MORE FROM GERALD BRACEY

BAIL ME OUT!
Handling Difficult Data and Tough Questions About Public Schools

Designed to make the reader more knowledgeable about the data concerning the achievements of American public schools. Sheds light on commonly held misconceptions about school performance.

Library binding, $65.95 (599.35 members) Trade paperback, $29.95 ($29.95 members)

The Truth About America’s Schools:
The Bracey Reports, 1991-97

Examines the myths, misconceptions, and misinformation about America’s schools. Includes seven Kaplan reports from the series beginning in 1991.

Trade paperback: $22 ($19.50 members)

Put to the Test: An Educator’s and Consumer’s Guide to Standardized Testing

A clear resource with objective information about what standardized tests can and cannot do, how they are constructed, and how they are used and misused.

Oversize paperback: $17.95 ($14.95 members) Buy 15 or more, only $9.55 each.

TO ORDER, PHONE 1-800-766-1156

Orr send institutional purchase order to Phi Delta Kappa International, P.O. Box 789, Bloomington, IN 47402-7769. Include shipping and handling charge on all orders: from $5.01 to $50, add $5; from $50.01 to $100, add $10; more than $100, add $15 plus 5% of merchandise total. Indiana residents add 5% sales tax. Purchase orders and credit card orders (VISA, MasterCard, American Express, and Discover) also are accepted by fax at 812/339-0018.

SOURCE CODE: S1182

OCTOBER 2001 123