

# **More Smoke and Mirrors: A Critique of the National Reading Panel (NRP) Report on "Fluency"**

Stephen Krashen

Phi Delta Kappan (October, 2001)

In her review of the National Reading Panel's (NRP) report on phonics, Elaine Garan concluded that the report involved "a limited number of studies of a narrow population. . ." <sup>1</sup> In this note, I will argue that this problem is not limited to the section on phonics: It also applies to the NRP's section 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 improves reading achievement. <sup>2</sup>

## **Omissions**

The selection criteria used by the NRP for selection of studies were as follows:

1. The study had to be a research study that appeared to consider the effect 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 appear in a refereed journal.
4. The study had to have been carried out with English language reading. <sup>3</sup>

The NRP claimed it could find only 14 studies that met these criteria. <sup>4</sup> Of these, 10 were studies of the impact of sustained silent reading (SSR) programs in which some classtime is set aside for free voluntary reading with little or no "accountability." Of these 10, three 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 other studies showed no difference or no gains. Table 1 summarizes these outcomes.

## **TABLE 1**

## Summary of National Reading Panel Results: Duration of Treatment and Outcomes<sup>5</sup>

Duration	Positive	No difference	Negative
Less than seven months	2	8	0
7 months--1 year	2	2	0
positive = students in sustained silent reading programs outperform comparisons Results include ten studies, 14 comparisons			

In other sections of the NRP report, such as the sections on phonics and phonemic 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 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 cases, the subjects were undergraduate college students. Subjects in some of the studies were students of English as a second language.<sup>6</sup> 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.

### TABLE 2

## Duration of Treatment and Outcomes of SSR Studies: Expanded Set<sup>7</sup>

<b>Duration</b>	<b>Positive</b>	<b>No difference</b>	<b>Negative</b>
Less than 7 months	7	13	3
7 months--1 year	9	11	0
Greater than 1 year	8	2	0

In the studies in table 2, SSR students did as well 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 out of ten 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 carried out by Isabel Schon, Kenneth Hopkins, and Carol Vojir, there was no difference between SSR students and comparison groups, but only five of the eleven SSR teachers actually carried out SSR conscientiously.<sup>8</sup> The classes taught by these five achieved significantly better gains.

In a study by Ruth Cline and George Kretke, another study showing no difference, subjects were junior high school students who were reading two years above grade level, and probably had already established a reading habit.<sup>9</sup> Similarly, in Zephaniah Davis' study of eighth graders, SSR helped medium level readers but not better readers.<sup>10</sup> 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 paper will improve 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 comparisons in 35 out of 36 comparisons. This suggests that the "violations" do not affect the central issue of whether encouraging recreational reading impacts literacy development. Even if one only allows 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. Carver and Liebert 's study<sup>11</sup> 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, 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 sustained silent reading programs allow access to any books readers want to read, do not use extrinsic motivators, do not make students accountable for what they read, provide a wide variety of books, and typically meet for a short time each day over a long period.<sup>12</sup>

The NRP claimed that the advantage shown by readers in Joanne Burley's study<sup>13</sup> was "small." Students in sustained silent reading 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 claims 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."<sup>14</sup> This is not accurate: Student assignment was in fact random and the four teachers were randomly assigned to one of the four groups.<sup>15</sup> 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<sup>16</sup> 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 1.005, 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."<sup>17</sup>

In Gary and Maryann Manning's study,<sup>18</sup> 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 each other, that is, when they discussed their reading with each other and shared books. The NPR refers to this group's advantage as "slight,"<sup>19</sup> but it is not clear how they 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 comparisons in reading comprehension.<sup>20</sup> 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 (pretest 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 reported that D. Ray Reutzel and Paul Hollingsworth found no differences between SSR and skills practice.<sup>21</sup> What the panel did not mention is that the entire treatment lasted only ten 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). Reutzel and Hollingsworth 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 of the learning and practicing of discrete comprehension skills."<sup>22</sup>

## **Additional Evidence**

It should also be pointed 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.<sup>23</sup> It is hard to attribute these gains to anything but reading. There are, in addition, compelling case histories that cannot be easily explained on the basis of 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."<sup>24</sup> Or consider the case of Ben Carson,<sup>25</sup> a neurosurgeon who says that his mother's insistence that he read two books a week (of his own choosing) 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, and it helped 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 have been done, "raise serious questions" about its efficacy.<sup>26</sup> 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 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 describing others incorrectly. Garan asks that we look beneath the smoke and behind the mirrors of the NRP phonics report. The same needs to be done with the report on "encouraging fluency."

## Notes

1. Elaine Garan, "Beyond the Smoke and Mirrors: A Critique of the National Reading Panel Report on Phonics," *Phi Delta Kappan*, vol. 82, 2001, pp. 500-506. Quote from p. 505.
2. The "fluency" report 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 <http://www.nichd.nih.gov/publications/nrp/smallbook.pdf>, 2000.
3. Ibid. p. 3-23.
4. The NRP report devoted only six pages to recreational reading. In contrast, sixty-six pages are devoted to phonemic awareness and nearly as many to phonics.

5. Studies included in Table 1 follow:

Duration less than seven months, positive:

Joanne Burley, "Short-term, High Intensity Reading Practice Methods for Upward Bound Students: An Appraisal." *Negro Educational Review*, vol. 31, 1980, pp 156-161; Judith Langford and Elizabeth Allen, "The Effects of U.S.S.R. on Students' Attitudes and Achievements." *Reading Horizons*, vol. 23, 1983, pp. 194-200;

Duration less than seven months, no difference:

Howard Evans and John Towner, "Sustained Silent Reading: Does it Increase Skills?" *Reading Teacher* vol.29, 1975, pp. 155-156; Cathy Collins, "Sustained Silent Reading Periods: Effects of Teachers' Behaviors and Students' Achievements." *Elementary School Journal*, vol. 81, 1980, pp. 109-114; Edward Summers and J. V. McClelland, "A Field-Based Evaluation of Sustained Silent Reading (SSR) in Intermediate Grades." *Alberta Journal of Educational Research*, vol. 28, 1982. pp. 110-112 (three groups); D. Ray Reutzel and Paul M. Hollingsworth, "Reading Comprehension Skills: Testing the Distinctiveness Hypothesis." *Reading Research and Instruction*, vol. 30, 1991, pp. 32-46; Sandra Holt and Frances O'Tuel. 1989. "The Effect of Sustained Silent Reading and Writing on Achievement and Attitudes of Seventh and Eighth Grade Students Reading Two Years Below Grade Level." *Reading Improvement*, vol. 26. 1989, pp. 290-297; Ronald Carver and Robert Liebert, "The Effect of Reading Library Books at Different Levels of Difficulty upon Gains in Reading," *Reading Research Quarterly*, vol. 30, 1995, pp. 26-48.

Duration seven months to one year, positive:

Zephaniah Davis, "A Comparison of the Effectiveness of Sustained Silent Reading and Directed Reading Activity on Students' Reading Achievement." *The High School Journal*, vol. 72, 1988, pp. 46-48; Gary Manning and Maryann Manning. 1984. "What Models of Recreational Reading Make a Difference?" *Reading World*, vol. 23, 1984, pp.375-380 (peer-interaction condition).

Duration seven months to one year, no difference:

Manning and Manning, op. cit. (pure SSR and student-teacher conference conditions).

6. In a response to a commentary of mine (Stephen Krashen, "Reading Report: One Research's Errors and Omissions," *Education Week*, May 10, 2000, p. 49), Timothy Shanahan, speaking for the NRP, notes that some studies were omitted because "The panel did not attempt to address second-language issues as the report clearly states." (see Timothy Shanahan, "Reading Panel: A Member Responds to a Critic," *Education Week*, May 31, 2000, p. 39). The NRP selection criteria do not mention English as a second language. Instead, the report notes only that studies had to be carried out with "English language reading" (p.3-23). Excluding second language acquirers resulted in ignoring some of the most important studies in the literature, both in terms of theory and practice.
7. Studies included in Table 2 follow.

Duration less than seven months, positive:

Burley, 1980, op.cit.; Anne Wolf and Larry Mikulecky, "Effects of Uninterrupted Sustained Silent Reading and of Reading Skills instruction on Changes in Secondary 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, Clemson, S.C., 1978) pp. 226-228.; Mabel Aranha, "Sustained Silent Reading Goes East," *Reading Teacher*, vol. 39, 1985, pp.214-217; Ira Gordon and Christine Clark. "An experiment in individualized reading," *Childhood Education* vol. 38, 1961, pp. 112-113; Holt and O'Tuel, op.cit.; Mary Husar, "Reading and more reading," *Elementary English* vol. 44, 1967, pp. 378-385, 382 (grade 6); Beniko Mason and Stephen. Krashen, "Extensive Reading in English as a Foreign Language," *System* vol.25, 1997, pp. 91-102 (study 1).

Duration less than seven months, no difference:

Evans and Towner, op.cit. ; Collins, op. cit.; Edith Sperzl., "The Effect of Comic Books on Vocabulary Growth and Reading Comprehension." *Elementary English*, vol. 25, 1948, pp. 109-113. Marvin Oliver, "The Effect of High Intensity Practice on Reading Comprehension." *Reading Improvement*, vol. 10, 1973, pp. 16-18; idem., "The Effect of High Intensity Practice on Reading Achievement," *Reading Improvement* vol.13, 1976, pp. 226-228.; Isabel Schon, Kenneth Hopkins, and Carol Vojir. "The Effects of Spanish Reading Emphasis on the English and Spanish Reading Abilities of Hispanic High School Students." *Bilingual Review* vol. 11, 1984, pp. 33-39 (Tempe group); Harry Sartain, "The Roseville Experiment with Individualized Reading." *Reading Teacher*, vol.12, 1960, pp. 277-281 ("good reader" group); Summers and McClelland, op. cit. (three groups); Husar, op. cit. (grades 4,5), Holt and O'Tuel, 1989, op. cit. (grade 8);

Duration less than seven months, negative: Hoyle Lawson, "Effects of Free Reading on the Reading Achievement of Sixth Grade Pupils," in J.A. Figuerel, ed. *Forging Ahead in Reading*, (Newark, Del: International Reading Association, 1968), pp. 501-504.; Sartain, op. cit. ("slow reader"



group); San Diego County, "A Plan for Research. In Sam Duker, ed. *Individualized Reading: Readings*,. (Metuchen, N.J. Scarecrow Press, 1965) pp. 359-363.

Duration seven months to one year, positive:

Warwick Elley, "Acquiring Literacy in a Second Language: The Effect of Book- Based Programs." *Language Learning*, vol.41, 1991, pp. 375-411 (Singapore, P1 survey); Marion Jenkins, "Self-Selection in Reading." *Reading Teacher* vol.11, 1957, pp. 84-90; Lois Bader, Jeannette. Veatch, and J. Lloyd Eldridge, "Trade Books or Basal Readers?" *Reading Improvement*, vol.24, 1987, pp. 62-67; Davis, op. cit. (medium ability readers); Mason and Krashen, op. cit. (studies 2a,2b, 3); Propitas Lituanas, George Jacobs, Willy. Renandya. "A Study of Extensive Reading with Remedial Reading Students," In Y. M. Cheah and S. M. Ng, eds. *Language Instructional Issues in Asian Classrooms* . (Newark: International Reading Association, 1999) pp. 89-104.

Duration seven months to one year, no difference:

Manning and Manning, op. cit. (pure SSR and student-teacher conference conditions); Schon, Hopkins and Vojir, 1984, op. cit. (Chandler School District); Isabel Schon, Kenneth Hopkins, and Carol Vojir, "The Effects of Special Reading Time in Spanish on the Reading Abilities and Attitudes of Hispanic Junior High School Students." *Journal of Psycholinguistic Research*, vol. 14, 1985, pp. 57-65 (grade 7, grade 8); James McDonald, Theodore Harris, and John Mann. "Individual Versus Group Instruction in First Grade Reading." *Reading Teacher*, vol.19, 1966, pp. 643-646, 652; Vincent Greaney, "A Comparison of Individualized and Basal Reader Approaches to Reading Instruction." *Irish Journal of Education*, vol.1, 1970, pp. 19-29; Floyd Davis and James Lucas, "An Experiment in Individualized Reading," *Reading Teacher*, vol. 24, 1971, pp. 737-743,747 (grades 7 and 8); Ann Healy, "Changing Children's Attitudes Toward Reading," *Elementary English*, vol.40, 1963, pp. 255-257, 279; Davis, op. cit. (high ability readers).

Duration longer than one year, positive:

Warwick Elley and Francis Mangubhai. 1983. "The Impact of Reading on Second Language Learning," *Reading Research Quarterly* vol.19, 1983. pp.53-67 (grades 4 and 5); Elley, op. cit. (Singapore, sample of 512, P3 survey), Miriam Aronow, "A Study of the Effect of Individualized Reading on Children's Reading Test Scores," *Reading Teacher*, vol. 15, 1961, pp. 86-91; Ben Bohnhorst and Sophia Sellars, "Individual Reading Instruction vs. Basal Textbook Instruction: Some Tentative Explorations," *Elementary English*, vol. 36, 1959, pp. 185-190, 202; Frances Cyrog, "Self-Selection in Reading: Report of a Longitudinal Study," In Malcolm Douglas, ed. *Claremont Reading*

*Conference: 26th Yearbook* (Claremont, Ca.: Claremont Graduate School, 1962) pp. 106-113; Rodney Johnson, "Individualized and basal primary reading programs," *Elementary English*, vol. 42, 1965, pp. 902-904, 915.

Duration longer than one year, no difference:

Ruth Cline and George Kretke, "An Evaluation of Long-Term SSR in the Junior High School, *Journal of Reading*, March, 1980, pp. 503-506; "Warwick Elley, Ian Barham, Hilary Lamb, and Malcolm Wyllie. 1976. "The Role of Grammar in a Secondary School Curriculum." *Research in the Teaching of English* vol.10, 1976. pp. 5-21.

Shanahan claimed that most omitted studies of sustained silent reading using native speakers of English were unpublished dissertations (Shanahan, op. cit.). None of the studies listed above are unpublished dissertations. All were published in refereed journals, except for two studies published in the National Reading Handbook Yearbook (Lawson, Wolf and Mikulecky), one from a book published by the International Reading Association (Lituanas, Jacobs, and Renandya), one study from the Claremont Conference Handbook (Cyrog), and one school district report (San Diego County).

8. Schon, Hopkins, and Vojir, "The Effects of Special Reading Time."
9. Cline and Kretke, op. cit.
10. Davis, op. cit.
11. Carver and Liebert, op. cit.
12. See e.g. Janice Pilgreen, *The SSR Handbook: How to Organize and Maintain a Sustained Silent Reading Program*. Portsmouth, NH: Heinemann, 2000.
13. Burley, op. cit.
14. Krashen, 2000, op. cit. and Shanahan, op. cit.
15. Burley, p. 158.
16. Langford and Allen, op. cit.
17. National Reading Panel, op. cit. p.3-26.
18. Manning and Manning, op. cit.
19. National Reading Panel, op. cit. p. 3-26
20. Holt and O'Teul, op. cit.; National Reading Panel, p. 3-25.
21. Reutzel and Hollingsworth, op. cit.; National Reading Panel, p. 3-24.
22. Reutzel and Hollingsworth, op. cit, p. 41.
23. For a review, see Stephen Krashen, *The Power of Reading*, (Englewood, CO: Libraries Unlimited, 1993).

24. Richard Wright, *Black Boy*, (New York: Harper and Row, 1966). p. 275.
25. Ben Carson, *Gifted Hands*, (Grand Rapids, MI: Zondervan Books, 1990).
26. National Reading Panel, *op. cit.*, p. 3-27.