# DOES READING MAKE YOU SMARTER? LITERACY AND THE DEVELOPMENT OF VERBAL INTELLIGENCE

*Keith E. Stanovich*

ONTARIO INSTITUTE FOR STUDIES IN EDUCATION
TORONTO, ONTARIO, CANADA M5S 1V6





ADVANCES IN CHILD DEVELOPMENT
AND BEHAVIOR, VOL. 24

## I. Introduction

What role do experiential differences play in determining variation in cognitive growth? This question has been at the heart of much theorizing in developmental psychology. Enthusiasm for experiential explanations of differences in cognitive growth has waxed and waned over the years, as has interest in explanations based on genetic inheritance. For example, interest in genetic determinants of differences in cognitive skills increased considerably in the 1980s (Plomin, DeFries, & McClearn, 1990; Thompson, Detterman, & Plomin, 1991), and some once-popular experiential hypotheses went out of favor altogether. Theories in which literacy is posited to be a determinate of individual differences in cognitive growth provide a case in point. Differential experience with print was once an important mechanism in many theories of cognitive change (Greenfield, 1972; Olson, 1977). In the 1980s, the idea that the acquisition of literacy has profound cognitive consequences went seriously out of favor (Erickson, 1984; Gee, 1988; Scribner & Cole, 1981; Street, 1984, 1988).

In this article, I argue that we should reconsider experience with print as an explanatory mechanism that can account for cognitive change. My argument begins with a review of selected literature on the consequences of literacy in which I claim that a role for reading experience in theories of cognitive change seems to have been prematurely dismissed. I then introduce a methodology for studying the cognitive consequences of literacy within a literate society that could help to revive research interest in this hypothesis.

## II. The Rise of the "Great Divide" Theories

Theories in which literate and nonliterate individuals and societies are posited to exhibit important cognitive differences have been termed *Great Divide theories* (Erickson, 1984; Olson & Torrance, 1991; Scribner & Cole, 1981). As is discussed later, in the 1980s the term was most often used pejoratively (e.g., Street, 1984). Great Divide theories can be partitioned according to whether they are concerned primarily with the indirect, mediated effects of literacy—habits of thought derived through cultural immersion in a literate society—or the direct, nonmediated effects of literacy on a particular individual's cognitive processes and knowledge structures (Goody, 1987; Scribner & Cole, 1978). For example, illiterates, or people who engage only marginally in literacy activities, may derive certain cognitive benefits from participation in a literate culture. These have been termed the mediated effects of literacy (Goody, 1987, pp. 217–252), and anthropologists and historians have done much work to assess these culturally mediated consequences of literacy. In contrast, psychologists have been

more concerned with examining the individual effects of having personally engaged in reading/writing activities (Scribner & Cole, 1978, 1981).

Great Divide theorists are further differentiated according to whether they posit a continuous range of effects, that is, effects that are linked to variation in print-related activities that exist even within literate populations. Greenfield's seminal (1972) work provides a case in point. Based on her cross-cultural research in Africa, Greenfield (1972) posited that facility with written language developed a set of cognitive competencies that were more elaborate than those associated with purely oral language. Her argument was based on the differences in context dependency between written and oral language:

> If a speaker of an oral language depends upon the surrounding context to communicate his message, then effective communication presupposes a common context and common point of view for both listener and speaker. The speaker, moreover, must assume that this is the case. He is, therefore, egocentric. . . . Speech based on a written language, in contrast, must be relatively independent of context for a number of reasons. (p. 170)

The central thesis of Greenfield's argument was that "Context-dependent speech is tied up with context-dependent thought, which in turn is the opposite of abstract thought" (p. 169). Greenfield outlined several examples of how context-independent language fosters abstract thought and problem solving. Although most of her discussion derived from her work done among the Wolof of Senegal, Greenfield (1972) extended her hypotheses to encompass cultural subgroups in the United States who were not illiterate but who had less exposure to written language. That is, she championed a continuous version of the Great Divide hypothesis by positing that differences in degrees of print exposure within even a generally literate society also have cognitive consequences.

Greenfield's hypotheses concerning the direct effects of literacy on an individual reader's cognition paralleled theories of the effects of literacy at the societal level that were popular among anthropologists and historians (Akinnaso, 1981; Goody, 1977, 1987; Havelock, 1963, 1980; Musgrove, 1982; Ong, 1967, 1982). Goody's influential writings (1977, 1980; Goody & Watt, 1968) contain hypotheses about the effects of literacy at the level of societies and cultures rather than individuals:

> The specific proposition is that writing, and more especially alphabetic literacy, made it possible to scrutinise discourse in a different kind of way by giving oral communication a semi-permanent form; this scrutiny favored the increase in scope of critical activity, and hence rationality, scepticism, and logic to resurrect memories of those questionable dichotomies. It increased the potentialities of criticism because writing laid out discourse before one's eyes in a different way; at the same time [it] increased the potentiality for cumulative knowledge, especially knowledge of an abstract kind, because it changed the nature of communication beyond that of face-to-face contact as well as the system for the storage of information; in this way a wider range of

> "thought" was made available to the reading public. . . . [it] enabled man to stand back from his creation and examine it in a more abstract, generalised, and rational way. (1977, p. 37)

Elaborations of this argument have echoed throughout the anthropological and historical literature on the effects of literacy (Havelock, 1963, 1980; Musgrove, 1982; Ong, 1982). Ong (1982) made the case for the cognitive effects of literacy most strongly: "Without writing, the literate mind would not and could not think as it does, not only when engaged in writing but normally even when it is composing its thoughts in oral form. More than any other single invention, writing has transformed human consciousness" (p. 78). The causal mechanism emphasized by Ong (1982) was a variant on the Goody/Greenfield theme: "Writing fosters abstractions that disengage knowledge from the arena where human beings struggle with one another. It separates the knower from the known" (pp. 43–44).

Olson (1977, 1986b, 1988) has presented a related causal theory of how literacy comes to influence thought. In his important 1977 essay Olson contrasted texts (written prose statements) with utterances (informal oral-language statements). His thesis, which he made clear was intended to apply to both the societal and individual consequences of literacy, was that "There is a transition from utterance to text both culturally and developmentally and this transition can be described as one of increasing explicitness, with language increasingly able to stand as an unambiguous or autonomous representation of meaning" (p. 258). Olson emphasized the importance of the assumption that meaning resides autonomously within the text and "the consequences of that assumption, particularly of the attempts to make it true" (p. 258). The highest form of the autonomous text ideal is the essayist technique: "The more fundamental effect of this approach to text was on the writer, whose task now was to create autonomous text—to write in such a manner that the sentence was an adequate, explicit representation of the meaning, relying on no implicit premises or personal interpretations" (p. 268). He argued that

> Logical development in a literate culture involves learning to apply logical operations to the sentence meaning rather than to the assimilated or interpreted or assumed speaker's meaning. Development consists of learning to confine interpretation to the meaning explicitly represented in the text and to draw inferences exclusively from that formal but restricted interpretation. . . . The developmental hypothesis offered here is that the ability to assign a meaning to the sentence per se, independent of its non-linguistic interpretive context, is achieved only well into the school years. (pp. 274–275)

Hypotheses of the type put forth by Greenfield, Goody, and Olson came to be termed Great Divide theories because of the number and importance of the cognitive consequences of literacy that were assumed. Among these were

> Logical and analytic modes of thought; general and abstract uses of language; critical and rational thought; a skeptical and questioning attitude; a distinction between myth and history; the recognition of the importance of time and space; complex and modern governments (with separation of church and state); political democracy and greater social equity; economic development. . . . It leads to people who are innovative, achievement oriented, productive, cosmopolitan, politically aware, more globally (nationally and internationally) oriented. (Gee, 1988, p. 196)

## III. Great Divide Theories under Attack

By the late 1970s and early 1980s the Great Divide position had garnered considerable support and momentum. By the late 1980s, however, the situation had changed drastically. By then, many scholars had accepted Graff's (1979, 1986, 1987) characterization of the Great Divide theory as the "literacy myth" and Gee (1988) was claiming that "At least in academic circles, the literacy myth is on its last legs" (p. 196). What happened to cause such a rapid and extreme theoretical reversal?

Not surprisingly, a number of interacting factors acted in concert to cause the collapse of the "literacy myth." First, literacy's effects at the societal level were brought into question by historical studies indicating that literacy was intertwined with certain cultural effects in a much more intricate and interactive way than was implied by some of the simpler theories that emphasized unidirectional causation. For example, the link between economic development and national levels of literacy has turned out to be much more complex than originally thought. Literacy levels are as much a consequence of economic development as they are its cause (Fuller, Edwards, & Gorman, 1987; Graff, 1986, 1987; Kaestle, 1991; Wagner, 1987).

The plausibility of literacy having an effect on cognition at the level of individuals has also been questioned. Some researchers have questioned the distinctions between utterance and text that served as the guiding assumptions of Great Divide theorizing. Clearly, for example, oral speech in formalized settings can contain all the features associated with written text: detachment, certain types of subordination, integration, nominalization (Biber, 1986; Chafe & Danielewicz, 1987; Feldman, 1991; Nystrand, 1987; Redeker, 1984; Tannen, 1982). Similarly, written texts need not always contain these features. Thus, utterances can sometimes have the characteristics of text and vice versa. Acceptance of the idea that the features of utterance and text that were allegedly responsible for differential cognitive effects were different only probabilistically rather than in a discrete sense had the effect of making the "divide" seem less "great."

Adding to these second thoughts about the consequences of literacy was a radical social critique that conceived of literacy as just one more mechanism used by powerful groups to maintain social privilege. For example, Street (1984)

claimed that "The actual examples of literacy in different societies that are available to us suggest that it is more often 'restrictive' and hegemonic, and concerned with instilling discipline and exercising social control" (p. 4), and that "Schooling and techniques of teaching literacy are often forms of hegemony" (p. 11). Another author titled her book *The Violence of Literacy* and argued that American society "stakes much on the oppressive powers of literacy" (Stuckey, 1991, p. 30) and that "it is possible that a system of ownership built on the ownership of literacy is more violent than past systems, however. Though it seems difficult to surpass the violence of systems of indenture, slavery, industrialism, and the exploitation of immigrant or migrant labor, literacy provides a unique bottleneck" (p. 18). Gee (1988) echoed the argument that "literacy has been used, in age after age, to solidify the social hierarchy, empower elites, and ensure that people lower in the hierarchy accept the values, norms and beliefs of the elites, even when it is not in their self-interest" (p. 205).

In short, the social effects of literacy are no longer universally viewed as positive, at least by some educational theorists. Great Divide theories seemed, to these same scholars, to be a case par excellence of blaming the victims. Thus, the popular social critiques of the 1980s were used as weapons against any theory of the consequences of literacy that posited substantial cognitive effects following from differential engagement in literacy activities. Because people *do* differentially engage in literacy activities, any such effects were bound to create cognitive inequalities that most neo-Marxist and socioconstructivist theorists were committed to denying. Great Divide theories ran straight into the brick wall of cultural and epistemological relativism that was a foundational assumption of these social critiques (for discussions, see Gellner, 1985; Hollis & Lukes, 1982; Musgrove, 1982, Shweder, 1991; Siegel, 1988; Sperber, 1985). The critiques seemed to rest on the tenuous assumption that literacy creates enormous sociopolitical differences that are not associated with any concomitant cognitive differences.

## IV. The Death Blow to Great Divide Theories: Scribner and Cole

Thus, a confluence of academic critiques in the 1980s undermined hypotheses about the effects of literacy; however, another factor was probably more influential than all the academic critiques combined. This factor was the investigation into literacy effects among the Vai in Africa by Scribner and Cole (1981).

The work of Scribner and Cole provides one of the firmest foundations for the assaults on the "literacy myth" and it is repeatedly cited in critiques of Great Divide theories:

> In the Scribner and Cole study, literacy in and of itself led to no grandiose cognitive abilities. (Gee, 1988, p. 203)
>
> The Vai findings caution us against such generalizations as are often made: that writing promotes general mental abilities. (Akinnaso, 1981, p. 175)
>
> Cole and Scribner point toward an interpretation that contradicts the usual view that literacy leads inevitably to higher forms of thought. . . . Such research may also control the assumptions and expectations that students carry to studies of literacy—such as presupposing literacy to be "liberating" or "revolutionary" in its consequences. There are, I suggest, better reasons to expect the opposite to be more often the case. (Graff, 1987, pp. 23–24)
>
> Scribner and Cole (1981) conclude that literacy per se contributes only marginally to cognitive development. (Nystrand, 1987, p. 236)

Scribner and Cole's investigation was ground breaking because they sought to separate the effects of literacy from the effects of schooling. The confounding of these two factors in earlier research (e.g., Greenfield, 1972; Luria, 1976) rendered tenuous any conclusions about the effects of literacy per se. Scribner and Cole took advantage of the fact that three scripts are in use among the Vai and that each script is associated with a particular context. English is learned in school and is used in formal settings (e.g., dealings with the government); Arabic is used for reading, writing, and memorizing the Koran; and an indigenous Vai script is transmitted outside of institutional settings and is used in personal correspondence and for some business purposes. Although some of the Vai are fluent in multiple scripts and some are illiterate, the fact that some individuals are familiar with only one script allows the separation of schooling effects from literacy effects. The comparison of Vai monoliterates with illiterates and individuals schooled in English is particularly diagnostic.

Scribner and Cole (1981) found no specific effect of Vai literacy on a number of tasks tapping general cognitive processes, including geometric sorting tasks, taxonomic categorization tasks, memory tasks, and syllogistic reasoning problems. Scribner and Cole concluded that "Effects of nonschooled literacies are spotty and appear on only a few performance measures. . . . These surely disappoint the grand expectations and lofty theories that inspired us to undertake this line of investigation" (pp. 130–132). The authors further argued that the findings "lay to rest some misconceptions about the psychology of literacy that went unchallenged in the past for lack of empirical data. First, it is clear from the evidence we reviewed that nonschooled literacy, as we found and tested it among the Vai, does not produce general cognitive effects as we have defined them. The small and selective nature of Vai script and Arabic influences on cognitive performance precludes any sweeping generalizations about literacy and cognitive change" (p. 132). Only when Scribner and Cole changed the focus of their research program to metalinguistic tasks more tightly and specifically linked to reading and writing (e.g., grammatical judgment, rebus reading, integrating syllables) did they find any specific effects of Vai literacy.

Despite the existence of some "spotty effects" (p. 244), Scribner and Cole's summary conclusion has been accepted by many investigators:

> Our results are in direct conflict with persistent claims that "deep psychological differences" divide literate and nonliterate populations. On no task—logic, abstraction, memory, communication—did we find all nonliterates performing at lower levels than all literates. Even on tasks closely related to script activities, such as reading or writing with pictures, some nonliterates did as well as those with school or literacy experiences. We can and do claim that literacy promotes skills among the Vai, but we cannot and do not claim that literacy is a necessary and sufficient condition for any of the skills assessed. (p. 251)

One indirect effect of the widespread acceptance of interpretations of the Scribner and Cole results was that in the 1980s very few cognitive and developmental psychologists conducted empirical studies on the individual consequences of literacy. The seeming conclusiveness of the Scribner and Cole investigation and the difficulty of conducting studies on these issues dampened enthusiasm for new empirical investigations of the effects of literacy. In the rest of this article, I reopen the issue, first by arguing that the consensus against the idea of profound cognitive consequences of literacy was arrived at too hastily and then by introducing a methodology for studying the cognitive consequences of literacy within a generally literate society.

## V. Premature Closure on the Consequences of Literacy?

Although acknowledging the ground-breaking nature of the Scribner and Cole project, I still would point out that their results are often overinterpreted in the literature on literacy. A major issue that is often glossed over—but that, interestingly, was raised by Scribner and Cole themselves—is whether the nature of Vai literacy was such that it provided a valid test of the claims of Great Divide theorists. Olson (1977), for example, was clear that the literacy on which he staked his claim is the high-level literacy characterized by the use of an essayist style. Goody (1987, p. 252), in his argument for the cultural effects of literacy, was also clear that he referred to the type of literacy that enables the reader to have access to a wide range of the world's accumulated knowledge. Vai literacy is simply not of this type. Individuals typically do not learn the Vai script until their late teens or twenties. It is used primarily for personal letter writing among people who know each other and for conducting business with those with whom one is familiar. Vai writing does not contain the autonomous essayist form that Olson (1977) argued is the causal mechanism that spurs cognitive change. Instead, the letter writing that looms so large in the society of Vai-script literates is highly personalized, assuming elaborate shared knowledge between the letter writer and the recipient (Scribner & Cole, 1981, pp. 71–75). Finally, the Vai

have no libraries in this script that can be used by individuals to access the world's storehouse of knowledge.

Scribner and Cole themselves admitted that "literate practices among the Vai are far more restricted than in technologically sophisticated societies. . . . Cultural heritage is transmitted orally in a way that does not depend upon texts. . . . It does not open doors to vicarious experience, new bodies of knowledge, or new ways of thinking about major life problems" (p. 238). They flatly conceded that "Vai script literacy does not fulfill the expectations of those social scientists who consider literacy a prime mover in social change" (p. 239). This admission is most often omitted in the discussions by authors who use Scribner and Cole's work to attack the "Literacy Myth." Theorists who refer to Scribner and Cole's work often put forth their interpretations without the qualifications introduced by the original authors—another case of prophets' disciples being more fanatical than the prophets themselves. In short, Scribner and Cole's research, influential and provocative though it was, should not be considered the final word on the issue of the cognitive consequences of literacy, although surely it would be foolish not to build on their insights.

## VI. Studying the Direct Consequences of Literacy within a Literate Society

### A. INTRODUCTION

Unfortunately, Scribner and Cole's innovative and costly project is unlikely to be replicated, so that resolving the issues using a variant of their methodology is not going to be possible; however, the cognitive consequences of literacy can be studied without necessarily using a cross-cultural comparison. I describe here a procedure for studying the cognitive consequences of literacy within a generally literate society. In developing the procedure, we exploited the fact that even within a generally literate culture, individuals vary tremendously in degree of exposure to print (Anderson, Wilson, & Fielding, 1988; Guthrie & Greaney, 1991; Guthrie & Seifert, 1983). Even among a group of individuals who have the same level of assessed reading comprehension ability, remarkably large differences are found in their degree of engagement in print-related activities (Stanovich & West, 1989) and the correlates of this natural variation can be studied. Comparing literates and illiterates is the exclusive design of choice only if the effects of literacy are believed to be completely discontinuous, with no cognitive consequences of variation in amount of print exposure among literate individuals. We speculate that the discontinuity assumption is false and that there is measurable cognitive variation among people who differ only in the amount of reading that they do.

In choosing which variables to focus on in our initial investigations, my research group was again influenced by the outcomes of Scribner and Cole's investigation. In a sense, we started where Scribner and Cole finished. That is, in the first part of their investigation, they concentrated on looking for effects of literacy on tasks that tapped developmental change in general cognitive processes. The tasks in the second part of their investigation—rebus reading, integrating auditory information, word pronunciation, and communication games—were more closely tied to aspects of Vai literacy and specific effects of literacy on these tasks were easier to demonstrate. In our research program on the cognitive consequences of differences in print exposure, we have inverted the investigative chronology of Scribner and Cole by starting with tasks that are more closely linked to literacy skills. Contingent on positive outcomes in these domains, we have examined more general cognitive processes. Thus, we established our methodology (see Stanovich & West, 1989) by examining criterion variables—orthographic knowledge and spelling—that should clearly be linked to individual differences in print exposure. We then expanded the set of criterion variables to encompass broader domains such as vocabulary, cultural knowledge, and verbal fluency.

### B. THE RESEARCH STRATEGY

In our methodology, we attempted to correlate differential engagement in reading activities with various cognitive outcomes that have been associated with the acquisition of literacy; however, such an experimental logic, if not supplemented with additional methodological controls, will yield data subject to an inordinately large number of alternative explanations. Levels of print exposure are correlated with too many other cognitive and behavioral characteristics. Avid readers tend to be different from nonreaders on a wide variety of cognitive skills, behavioral habits, and background variables (Guthrie, Schafer, & Hutchinson, 1991; Kaestle, 1991; Zill & Winglee, 1990). Attributing any particular outcome to print exposure uniquely is thus extremely difficult.

We have used a hierarchical regression logic first introduced by Anderson et al. (1988) to deal with this problem. The logic of the regression analysis allows any control variables entered first into the regression equation to explain any variance that they can in the criterion variable. Following these control variables, the measures of print exposure are entered. Thus, the procedure allows the investigator to assess whether reliable variance remains to be explained after the control variables are entered and whether print exposure is associated with this remaining variance. In our analyses, we first regressed out ability measures most likely to lead to spurious relationships before examining the linkage between print exposure and criterion variables.

This procedure of reducing possible spurious relationships by first partialing

relevant ability measures was used in our early investigations of subword processes in reading. For example, in previous work we had demonstrated that, independently of decoding ability, variation in print exposure among adults predicts variation in specific types of orthographic knowledge (Stanovich & West, 1989). Similarly, in a study of children's performance (Cunningham & Stanovich, 1990) we found that after partialing out IQ, memory ability, and phonological processing abilities, print exposure accounted for significant variance in orthographic knowledge and word recognition. The logic of our analytic strategy is quite conservative because in certain analyses we have actually partialed out variance in abilities that are likely to be developed by print exposure itself (Stanovich, 1986); however, the explanatory ambiguities surrounding a variable such as print exposure have led us to continue to structure the analyses in a "worst case" manner, as far as print exposure is concerned.

### C. ASSESSING PRINT EXPOSURE: THE DIARY TECHNIQUE

A variety of methods have been used to assess individual differences in exposure to print. For example, a variety of questionnaire and interview techniques have been used to assess relative differences in print exposure (e.g., Estes, 1971; Guthrie, 1981; Guthrie & Greaney, 1991; Guthrie & Seifert, 1983; Lewis & Teale, 1980; Sharon, 1973–1974; Walberg & Tsai, 1984), but many of these are encumbered with reliability and validity problems. A more valid method, but also a more logistically complicated one, is the use of daily activity diaries filled out by subjects (Anderson et al., 1988; Greaney, 1980; Greaney & Hegarty, 1987; Rice, 1986; Taylor, Frye, & Maruyama, 1990). Activity diaries yield estimates of the actual amount of time spent on literacy activities and are generally more valid than interview or questionnaire instruments (Carp & Carp, 1981).

Anderson et al. (1988) used the activity diary method to estimate the amount of time that fifth graders (10- to 11-year-olds) spent reading in their nonschool hours. They found that time spent reading predicted fifth-grade reading comprehension after the variance in second-grade (7- to 8-year-olds) reading comprehension had been controlled. This result seems to indicate that exposure to print was a contributor to individual differences in growth in reading ability over the elementary school years. My research group has completed a series of investigations in which we attempted to determine whether the specific correlates of print exposure can be extended beyond the demonstration by Anderson et al. (1988). We have employed the activity diary method in some of our own studies.

Our method of collecting daily activity records was adapted from the nonschool time diary investigation of Anderson et al. (1988), but we also attempted to improve on their methods in several respects (see Allen, Cipielewski, & Stanovich, 1992). Our daily activity record-keeping procedure was designed to minimize the time students would need to spend on it; to minimize the necessity

for adding and subtracting minutes or converting hours into minutes; and to maximize student judgment accuracy. We collected data over a 3-week period and thus obtained estimates of the average number of minutes per day that the children in our fifth-grade (10- to 11-year-olds) sample spent in various activities when they were outside of school.

Table I shows the mean and median minutes per day spent during the non-school hours in the various categories of activity that were listed on the activity records (with the exception of all reading, which is a composite of the categories book reading, comic reading, and other reading). The figure for book reading in the table includes only those instances in which the child could give at least a fragment of the title or a character from the book (a procedure adapted from Anderson et al., 1988). In contrast, all reading, the most liberal category, included all instances where book reading was indicated on the diary sheets.

That the means were generally larger than the medians reflects the positive skew of most of the variables, particularly the reading variables. The positive skew of reading time has repeatedly been observed in activity diary studies (Anderson et al., 1988; Greaney, 1980). Although some of our categories were different from those of the Anderson et al. (1988) study, those that were common

TABLE I

Time Spent per Day in Various Activities by a Fifth-Grade Sample[a]

| Activity | Minutes per day[b] | | |
|---|---|---|---|
| | Mean | Median | *SD* |
| Reading | 21.3 | 16.0 | 19.4 |
| Books | 10.2 | 5.0 | 15.0 |
| Comics | 2.1 | 0.0 | 4.0 |
| Other | 5.8 | 3.5 | 8.4 |
| Television watching | 83.2 | 68.0 | 65.5 |
| Eating | 52.9 | 54.6 | 19.1 |
| Homework | 49.0 | 45.0 | 26.9 |
| Just playing around | 35.7 | 26.0 | 32.7 |
| Playing outdoor games | 25.7 | 18.0 | 25.2 |
| Talking | 17.0 | 13.6 | 12.9 |
| Family activities | 18.0 | 13.9 | 18.2 |
| Playing indoor games | 14.5 | 8.6 | 19.1 |
| Practices | 14.3 | 8.0 | 18.4 |
| Hobbies | 7.9 | 0.0 | 17.5 |
| Chores | 6.6 | 5.0 | 6.4 |
| Lessons | 4.5 | 0.0 | 7.0 |
| Other | 62.6 | 60.0 | 35.0 |

[a] Adapted from Allen, Cipielewski, and Stanovich (1992).

[b] Weekdays and weekends are proportionately represented.

were ordered similarly in the two studies. For example, television watching was the most frequent activity and book reading was far down the list in both studies. Our fifth graders watched less television (83.2 minutes versus 131.1 minutes) and did more homework (49.0 minutes versus 18.9 minutes) than the Anderson et al. fifth graders. These differences might reflect the use of different populations—a private school in our study and public schools in the Anderson et al. (1988) study. Previous studies have shown private/public school differences in television and homework habits (Coleman, Hoffer, & Kilgore, 1982).

Despite differences in the estimates in other categories, our estimates of book reading time (mean and median of 10.2 and 5.0 minutes, respectively) are very close to those obtained in the Anderson et al. study (10.1 and 4.6 minutes). Certain rough generalizations thus hold across the two studies: Fifth graders (10- to 11-year-olds) spend around 5 minutes per night reading books for pleasure outside of school, roughly one-tenth the amount of time they spend watching television. These figures call to mind the many studies of school achievement in which American children scored poorly and in which their poor performance was linked to excessive television watching, low levels of homework, and little reading (Applebee, Langer, & Mullis, 1988; Chen & Stevenson, 1989; Stevenson et al., 1985).

Our specific concern, however, was whether children's reading volume related to their achievement and whether such a linkage could be shown to have any specificity. Book reading time (logarithmically transformed, see Allen et al., 1992, and Anderson et al., 1988) correlated .39 with a standardized test of vocabulary knowledge. We attempted a further assessment of the specificity of the relation between book reading and vocabulary development by conducting a hierarchical regression analysis in which the standardized vocabulary test was the criterion measure and in which performance on a standardized mathematics test was forced into the equation first as a control for general school learning ability. When entered second, book reading time explained an additional 9.7% of the variance and this unique variance explained was statistically significant ($p < .01$). Thus, the linkage between vocabulary and book reading time remains even when variability in general academic performance is partialed out.

### D. ASSESSING PRINT EXPOSURE: THE RECOGNITION CHECKLIST TECHNIQUE

In the study described in Section VI,C we employed a comprehensive activity recording methodology in which children accounted for all of their out-of-school time over a period of 3 weeks. This methodology provides not only estimates of relative differences in print exposure among children, but also estimates of the actual amount of time (in minutes per day) spent on literacy activities. The measurement of absolute amounts of reading activity and the methodologies used

to achieve such measurement have, however, a number of associated problems. First, the daily activity diary methodology requires extensive cooperation from teachers and students. Children must record, either at the end of the day or on the following morning, their activities from the day before and these recordings must be checked by a teacher or other adult to ensure that the scale is being used properly. Such a level of participant involvement may discourage many investigators from using the technique.

An additional problem is that the retrospective estimation of periods of time is a notoriously difficult task, even for adults (Bradburn, Rips, & Shevell, 1987; Burt & Kemp, 1991). This difficulty places some limits on how valid such estimates can be, even for a group of conscientious and well-motivated children. Finally, social desirability is a potential confound: Responses may be distorted because of tendencies to overreport socially desirable behaviors (Furnham, 1986; Paulhus, 1984)—in this case, to report more reading than actually takes place. Independent evidence indicates that social desirability does distort self-reports of book reading by adults (Ennis, 1965; Sharon, 1973–1974; Zill & Winglee, 1990). The extent to which it is a factor in children's self-reports of reading time is unknown.

The correlates of differential exposure to print can, however, be studied without estimating absolute amounts of reading minutes per day. Only an index of relative differences in exposure to print is required. Thus, one can use measures of print exposure that do not have some of the drawbacks of the activity diary method. My research group (Cunningham & Stanovich, 1990, 1991; Stanovich & West, 1989) has attempted to develop and validate measures of individual differences in print exposure that were designed (1) to yield estimates of relative differences in print exposure in a single 5- to 10-minute session, (2) to have very simple cognitive requirements (i.e., not requiring retrospective time estimates), and (3) to be immune from contamination from the tendency to give socially desirable responses.

The first measures we developed were designed for use with adult subjects. The Author Recognition Test (ART) and the Magazine Recognition Test (MRT) both exploited a signal detection logic whereby actual target items (real authors and real magazines) were embedded among foils (names that were not authors or magazine titles, respectively). Subjects simply scan the list and check the names they know to be authors on the ART and the titles they know to be magazines on the MRT. The measures thus have a signal detection logic. The number of correct items checked can be corrected for differential response biases that are revealed by the checking of foils. Although checklist procedures have been used before to assess print exposure (Chomsky, 1972), our procedure is unique in using foils to control for differential response criteria (see Stanovich & Cunningham, 1992, for examples of the stimuli).

In constructing the list of ART authors, items were selected who were most

likely to be encountered outside the classroom, so that the ART would be a proxy measure of out-of-school print exposure rather than of curriculum exposure. Thus, an attempt was made to avoid authors who are regularly studied in the school curriculum. For example, none of the authors that we have employed appeared in Ravitch and Finn's (1987) survey of the high-school literature curriculum. In short, the ART was intentionally biased toward out-of-school reading, because it was intended as an indirect measure of amount of free reading. The ART is dominated by "popular" authors. That is, it is not composed of "highbrow" writers who would be known by only the most highly educated or academically inclined readers. Instead, many of the book authors regularly appear on best seller lists and most have sold hundreds of thousands, if not millions, of copies. Although no statistical sampling of authors was carried out, an attempt was made to mix writers from a wide variety of genres.

Similarly, the sampling of titles on the MRT was deliberately biased toward popular publications. "Highbrow," academic, and low-circulation small-press publications that would be known by only the most highly educated or academically inclined readers were avoided. The publications on the MRT almost all have circulations in the hundreds of thousands, in many cases, millions. The foil names on the MRT do not appear in the 60,000 listings in *The Standard Periodical Directory* (Manning, 1988).

This checklist method has several advantages. First, it is immune to the social desirability effects that may contaminate responses to subjective self-estimates of socially valued activities such as reading. Guessing is not an advantageous strategy because it is easily detected and corrected for by an examination of the number of foils checked. Further, the cognitive demands of the task are quite low. The task does not necessitate frequency judgments, as do most questionnaire measures of print exposure, nor does it require retrospective time judgments, as does the use of daily activity diaries. Finally, the measures can be administered in a matter of a few minutes.

These checklist tasks are of course proxy indicators of a person's print exposure rather than measures of absolute amounts of reading in terms of minutes or estimated words (Anderson et al., 1988). The fact that the measures are very indirect proxy indicators is problematic in some contexts, but it is also sometimes a strength. Clearly, hearing about a magazine or author on television without having been exposed to the actual written work is problematic. The occurrence of this type of situation obviously reduces the validity of the tasks; however, a postexperimental comment sometimes made by adult subjects in our studies is worth noting: Some subjects said they knew a certain name was that of an author, but had never read anything that the author had written. When questioned about how they knew that the name was a writer, the subjects often replied that they had seen one of the author's books in a bookstore, had seen an author's book in the "new fiction" section at the library, had read a review of the author's work in

*Newsweek,* had seen an advertisement in the newspaper, and so on. In short, knowledge of that author's name was a proxy for reading activities, despite the fact that the particular author had not actually been read. Thus, although some ways of gaining familiarity with author names would reduce validity (TV, radio), most behaviors leading to familiarity with the author names are probably reflections of immersion in a literate environment.

We have developed analogous checklist measures for assessing children's exposure to print. One task is the Title Recognition Test (TRT), a measure that has the same signal detection logic as the adult ART and MRT, but involves children's book titles rather than authors as items. This children's measure shares the same advantages of immunity from socially desirable responding, objective assessment of response bias, low cognitive load, and lack of necessity for retrospective time judgments. The TRT consists of an intermixed list of actual children's book titles and foils for book names (see Allen et al., 1992, and Cunningham & Stanovich, 1991, for examples of stimuli). The titles used were selected from a sample of book titles generated in pilot investigations by groups of children ranging in age from second grade (7 to 8 years old) through high school, by examining various lists of children's titles, and by consulting teachers and reading education professionals knowledgeable about current trends in children's literature. In selecting the items to appear on the TRTs used in our investigations, we attempted to choose titles that were not prominent parts of classroom reading activities in the schools in which our studies were to be conducted. Because we wanted the TRT to reflect out-of-school rather than school-directed reading, we attempted to avoid books that were used in the school curriculum. Thus, if the test is used for this purpose, versions of it will necessarily differ somewhat in item content from classroom to classroom and from school to school.

To complement the TRT, we have also developed children's versions of the ART. Although it had originally been felt that the use of authors in a recognition checklist might be too difficult for children, pilot work in several classrooms indicated that children of this age could successfully respond to an author recognition measure and that performance on the task was diagnostic even though, as expected, recognition performance on the ART tends to be lower than that on the TRT. Authors on the measure were chosen using the same procedures employed for the TRT (see Allen et al., 1992). The score on all of these checklists—both child and adult versions—was the proportion of correct items checked minus the proportion of foils checked. This is the discrimination index from the two-high-threshold model of recognition performance (Snodgrass & Corwin, 1988). Other corrections for guessing and differential criterion effects (Snodgrass & Corwin, 1988) produce virtually identical correlational results.

Although the checklist measures have some obvious drawbacks as indices of children's exposure to print and degree of immersion in a literate environment,

just how much their obvious limitations impair their performance as probes of environmental print exposure is not known. For example, to get credit for a correct item on the TRT one clearly need have only some familiarity with the title. One need not have read the entire book or have remembered any of the contents at all. This seemingly problematic feature—that responses can be based on general familiarity rather than a more complete reading of the book—may be a strength just as often as a drawback. The possibility of responding on the basis of a shallow familiarity means that the TRT is not cognitively demanding and that it does not stress memory as much as some other tasks in which children might be asked to recall titles or information about plot and/or characters. The response demands of such tasks would necessarily implicate name retrieval and memory processes of considerable complexity (Bradburn et al., 1987; Burt & Kemp, 1991) that may affect performance and make such measures weaker indices of print exposure. Also, requiring recall of children may fail to index books read so long ago that they are partially forgotten. Title recognition appropriately allows such imperfectly recalled items to influence the obtained print exposure score. Anderson and Freebody (1983) reiterated all these arguments about task complexity when they argued for the relative purity of yes/no detection-type vocabulary measures over other indicators (see also Cooksey & Freebody, 1987; Meara & Buxton, 1987; White, Slater, & Graves, 1989).

### E. VALIDATING CHECKLIST MEASURES OF PRINT EXPOSURE

Because Anderson et al. (1988) have established the reliability and validity of the activity diary method of estimating print exposure, their methodology might well be considered the canonical method for assessing print exposure. Thus, we have used it as a baseline for assessing other methods. In an attempt to see whether questionnaire and recognition checklist measures of print exposure were measuring the same construct as the home reading-time estimates from children daily activity diaries, we included all these methods in a study of 57 fifth-grade (10- to 11-year-olds) children (Allen et al., 1992).

Table II shows a correlation matrix of all the media exposure measures used in the study. The variables are (1) book reading minutes as estimated from the activity diary, (2 and 3) two versions of the TRT, (4) one version of the ART, (5) a comics recognition checklist instrument modeled on the TRT, (6) number of preferences for reading on a questionnaire structured around forced choice between activities, (7) a reading disposition item from a typical reading habits questionnaire, (8) the recreational reading and (9) academic reading scales from the Elementary Reading Attitude Survey (ERAS, a public-domain questionnaire designed to give teachers an easy method of assessing attitudes toward reading; see McKenna & Kear, 1990), (10) the diary estimate of minutes of television watching each night, (11) number of choices of television on the activity prefer-

TABLE II

Intercorrelations of Differing Measures of Reading Habits[a]

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Book reading (diary) | | | | | | | | | | | |
| 2. TRT—form 1 | .48* | | | | | | | | | | |
| 3. TRT—form 2 | .43* | .65* | | | | | | | | | |
| 4. ART | .52* | .70* | .52* | | | | | | | | |
| 5. Comics recognition | .11 | .38* | .27 | .35* | | | | | | | |
| 6. Activity preference—reading | .25 | .34* | .34* | .35* | .40* | | | | | | |
| 7. Reading disposition question | .41* | .56* | .54* | .49* | .31* | .47* | | | | | |
| 8. ERAS—recreational | .37* | .39* | .37* | .34* | .24 | .55* | .54* | | | | |
| 9. ERAS—academic | −.08 | −.05 | .12 | −.05 | .03 | .30* | .10 | .52* | | | |
| 10. Television (diary) | −.32* | −.20 | −.17 | −.26 | .14 | .03 | −.18 | −.10 | .06 | | |
| 11. Activity preference—TV | −.13 | −.09 | −.04 | −.27* | −.11 | −.40* | −.22 | −.02 | −.09 | .33* | |
| 12. Television composite | −.22 | −.22 | −.19 | −.18 | −.01 | −.07 | −.32* | .06 | .26 | .63* | .44* |

[a] Correlations involving book reading (diary), TRT (form 1), ART, activity preference—reading, reading disposition question, and all television measures are based on an $N$ of 57. Correlations involving comics recognition are based on an $N$ of 53. Correlations involving TRT (form 2) and the ERAS are based on an $N$ of 43. The asterisk indicates correlations significant at the .05 level (two-tailed).

ence scale, and (12) a composite of television items from a typical media habits questionnaire.

Generally, the print exposure measures had significant correlations with each other and the television exposure measures were significantly intercorrelated. The print and television measures did not correlate with each other. In fact, most of these correlations were negative, although many did not attain statistical significance. This pattern of correlations suggests that the measures have both convergent and discriminant validity.

In Table II we can examine more closely the question of whether the time estimates of reading activity derived from the diary method correlate with the recognition checklist measures we have developed. As can be seen in rows 1 to 4, book reading time as measured by the activity diaries correlated significantly with both forms of the TRT and with the ART. These correlations are, in fact, substantial when we consider that these tasks had modest reliabilities (ranging from .68 to .86). Corrected for attenuation (Ghiselli, Campbell, & Zedeck, 1981, p. 241), the correlations of book reading minutes from the activity diaries with the two forms of the TRT are .65 and .59, and with the ART, .70. Additionally, the correlation between the two forms of the TRT was acceptably high as a parallel-forms reliability coefficient, and the correlations of these forms with the ART were acceptably high as concurrent validity estimates. Collectively, these results indicate that the recognition checklist measures are tapping a common construct with book reading minutes as estimated from the daily activity records. We would argue, quite simply, that the construct is exposure to print outside of school.

Book reading minutes from the activity diaries did not, however, correlate with the comics recognition measure, and the comics recognition measure correlated only modestly with the TRT and ART. This finding is consistent with previous results indicating that comic book reading is, functionally, somewhat different from book reading (Anderson et al., 1988; Greaney, 1980).

As a further method of exploring the relationships among the reading habits and attitudes measures, the variables listed in Table II were subjected to several methods of factor analysis. Different techniques of commonality estimation (including principal-components solutions) were tried, and several orthogonal and oblique rotations were computed. The results of a typical solution are displayed in Table III. In this analysis, squared multiple correlations were used as commonality estimates, iteration to a stable solution was carried out, and a varimax rotation was employed. The three factors retained accounted for 52.9% of the total variance. Table III lists all factor loadings greater than .250.

Factor 1 is clearly the most general print exposure factor, receiving loadings from all reading habits and attitudes measures except the ERAS academic scale. Factor 2 is clearly a television habits and attitudes factor. Factor 3 loads on questionnaire measures of attitudes about reading rather than indicators of actual behaviors (like the activity diary and recognition checklist measures). This pattern, in conjunction with the pattern displayed for Factor 1, suggests that Factor 3 reflects "disposition toward reading"—an attitudinal index of subjective feelings about the activity of reading. Factor 1, in contrast, might be interpreted as a latent index of actual print exposure. Although some of the questionnaire measures load on it, this factor is dominated by the print recognition measures and it

TABLE III

Factor Loadings for All Measures after Varimax Rotation[a]

| Measure | Factor | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1. Book reading (diary) | .629 | — | — |
| 2. TRT—form 1 | .870 | — | — |
| 3. TRT—form 2 | .661 | — | — |
| 4. ART | .768 | — | — |
| 5. Comics recognition | .419 | — | — |
| 6. Activity preference—reading | .442 | — | .596 |
| 7. Reading disposition question | .674 | — | .292 |
| 8. ERAS—recreational | .437 | — | .660 |
| 9. ERAS—academic | — | — | .697 |
| 10. Television (diary) | — | .669 | — |
| 11. Activity preference—television | — | .423 | — |
| 12. Television composite | — | .900 | — |

[a] Factor loadings lower than .250 have been eliminated.

also loads with the book reading minutes estimate from the activity diary. The fact that the recognition checklist measures loaded with the diary measure increases our confidence that the recognition checklist measures are a converging index of print exposure in children.

The checklist measures were also analyzed with the regression logic described earlier, to examine whether these print exposure measures can account for variance in vocabulary development after scores on a mathematics achievement test had been entered into the regression equation. The results for each of the four recognition checklist measures are present in Table IV. The outcome at the second step indicates whether each recognition checklist measure can predict vocabulary scores after controlling for general learning ability in a largely nonverbal domain. Three of four checklists (TRT 2, ART, and comics recognition) could do so. Although form 1 of the TRT predicted an additional 6.3% of the variance in vocabulary scores after mathematics computation scores had been entered into the equation, this estimate of the amount of unique variance explained did not reach statistical significance ($.05 < p < .10$).

A further test of the convergent validity of the checklist measures is provided by examining step 3 of the regressions, where the diary book reading time

TABLE IV

Hierarchical Regressions Predicting Vocabulary Scores[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| 1. Mathematics subtest | .407 | .166 | .166 | 10.93** |
| 2. TRT—form 1 | .472 | .223 | .063 | 3.96 |
| 3. Book reading (diary) | .524 | .275 | .052 | 3.80 |
| 1. Mathematics subtest | .274 | .075 | .075 | 3.57 |
| 2. TRT—form 2 | .432 | .187 | .112 | 5.93* |
| 3. Book reading (diary) | .462 | .213 | .026 | 1.41 |
| 1. Mathematics subtest | .407 | .166 | .166 | 10.93** |
| 2. ART | .583 | .340 | .174 | 14.24** |
| 3. Book reading (diary) | .595 | .354 | .014 | 1.18 |
| 1. Mathematics subtest | .400 | .160 | .160 | 10.27** |
| 2. Comics recognition | .565 | .319 | .159 | 12.39** |
| 3. Book reading (diary) | .636 | .405 | .086 | 7.52** |

[a] The first and third regressions are based on an $N$ of 57, the second regression is based on an $N$ of 46, and the fourth is based on an $N$ of 56.

*$p < .05$. **$p < .01$.

estimates were forced into the equation as the third variable. Investigating whether the diary estimates can predict variance after the checklist measures have been entered addresses the question of whether the vocabulary variance explained by the checklist measures overlapped with that explained by the book reading measure from the diary. In three of the four regressions, once the checklist measure was entered, book minutes from the diary no longer independently predicted vocabulary score (this outcome also occurred even when mathematics achievement scores were not entered into the equation). Thus, the variance in vocabulary scores explained by book reading minutes is variance that is largely shared with the checklist print exposure measures. The only case in which book reading minutes predicted vocabulary scores after the mathematics test and a checklist measure were in the equation was in the regression involving comics recognition. This result is consistent with the assumption that the comics recognition measure is a proxy for comic reading specifically and not generic literacy activity. This outcome supports the argument of previous investigators that comic reading has cognitive correlates different from book reading (e.g., Greaney, 1980).

## VII. The Specific Cognitive Correlates of Print Exposure

### A. PRINT EXPOSURE AS A CONTRIBUTOR TO GROWTH IN COMPREHENSION ABILITY

The preceding section indicates that the recognition checklists have convergent and discriminant validity as measures of print exposure. They therefore facilitate investigation of the specific cognitive correlates of exposure to print because they provide an alternative to the logistically difficult activity diary technique. Thus, my research group has embarked on a series of studies in which we employed recognition checklist measures and the regression logic outlined previously to see whether print exposure is a specific predictor (that is, after various ability controls are employed) of a variety of verbal skills.

First, we asked whether the recognition checklist measures of print exposure predict growth in reading ability throughout the elementary school years, as did the diary estimate of book reading time employed by Anderson et al. (1988). The regression analyses presented in Table V were addressed to this issue. They display the results of a study (Cipielewski & Stanovich, 1992) in which growth in reading comprehension ability was tracked by administering the comprehension tests from the Stanford Diagnostic Reading Tests and Iowa Tests of Basic Skills (ITBS) to 82 fifth graders who had been administered the comprehension subtest from the ITBS in the third grade (8- to 9-year-olds). The regressions are

TABLE V

Hierarchical Regressions Predicting Fifth-Grade Reading Ability[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| *Fifth-Grade Stanford Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .645 | .416 | .416 | 54.06** |
| 2. Title Recognition Test | .725 | .526 | .110 | 17.38** |
| *Fifth-Grade Stanford Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .591 | .349 | .349 | 34.89** |
| 2. Author Recognition Test | .655 | .430 | .081 | 9.02** |
| *Fifth-Grade Iowa Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .545 | .297 | .297 | 33.78** |
| 2. Title Recognition Test | .609 | .371 | .074 | 9.25** |
| *Fifth-Grade Iowa Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .485 | .236 | .236 | 20.95** |
| 2. Author Recognition Test | .503 | .253 | .017 | 1.56 |

[a] The italic spanner headings identify the dependent variables in the regression analyses.
**$p < .01$.

hierarchical forced-entry analyses for prediction of fifth-grade reading comprehension ability. Third-grade reading comprehension was entered first, followed by a recognition checklist measure of print exposure (either a version of the TRT or the ART). Thus, the analyses are essentially addressed to the question of whether the indicators of exposure to print can predict individual differences in growth in reading comprehension from third grade to fifth grade.

In three of four cases, print exposure measures predicted variance in fifth-grade reading comprehension ability after third-grade reading comprehension scores had been partialed out. Both the TRT and ART explained unique variance in fifth-grade Stanford comprehension scores, and the TRT (but not the ART) explained unique variance in fifth-grade Iowa comprehension scores. Table VI shows similar analyses of the sixth-grade reading comprehension performance of a smaller group of children ($N = 31$). Three of the four analyses demonstrated that print exposure measures predict individual differences in third- to sixth-grade growth in reading ability. The TRT predicted growth in reading comprehension when the criterion was either the Stanford or Iowa test; the ART predicted individual differences in comprehension when the Iowa test was the criterion but not when the Stanford test was the criterion. For the Stanford test, the unique variance explained (9.6%) was comparable to that observed in the analyses in

TABLE VI

Hierarchical Regressions Predicting Sixth-Grade Reading Ability[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| *Sixth-Grade Stanford Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .548 | .300 | .300 | 11.13** |
| 2. Title Recognition Test | .630 | .396 | .096 | 4.00* |
| *Sixth-Grade Stanford Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .469 | .220 | .220 | 6.20* |
| 2. Author Recognition Test | .562 | .316 | .096 | 2.96 |
| *Sixth-Grade Iowa Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .617 | .380 | .380 | 17.80** |
| 2. Title Recognition Test | .712 | .506 | .126 | 7.15* |
| *Sixth-Grade Iowa Reading Comprehension* | | | | |
| 1. Iowa Comprehension (Third) | .543 | .295 | .295 | 10.46** |
| 2. Author Recognition Test | .652 | .425 | .130 | 5.43* |

[a] The italic spanner headings identify the dependent variables in the regression analyses.
*$p < .05$. **$p < .01$.

Table V (8.1%), but it did not attain significance because of the smaller sample size.

In summary, to a large extent our work with the recognition checklist measures confirmed the finding of Anderson et al. (1988) with the activity diary that individual differences in exposure to print explain much of the variance in the development of comprehension abilities.

### B. PRINT EXPOSURE AS A CONTRIBUTOR TO GROWTH IN OTHER VERBAL SKILLS

In several studies, we asked whether print exposure contributes to growth in other cognitive skills. There are a number of reasons for expecting that the answer to this question might be affirmative. The study described in the preceding subsection indicated a unique contribution of print exposure to the explanation of reading comprehension, and reading comprehension is an extremely broad skill. A large body of research has demonstrated that reading skill is linked to a wide range of verbal abilities: Vocabulary, syntactic knowledge, metalinguistic awareness, verbal short-term memory, phonological awareness, speech production, inferential comprehension, semantic memory, and verbal fluency

form only a partial list (Byrne, 1981; Carr & Levy, 1990; Chall, 1983; Cunningham, Stanovich, & Wilson, 1990; Curtis, 1980; Just & Carpenter, 1987; Kamhi & Catts, 1989; Oakhill & Garnham, 1988; Palmer, MacLeod, Hunt, & Davidson, 1985; Perfetti, 1985; Siegel & Ryan, 1988, 1989; Stanovich & Cunningham, 1991; Stanovich, Cunningham, & Feeman, 1984; Stanovich, Nathan, & Zolman, 1988; Vellutino & Scanlon, 1987).

In certain domains, reading is especially likely to be a substantial contributor to cognitive growth. For example, as a mechanism for building content knowledge structures (Glaser, 1984), reading seems to be unparalleled (Goody, 1987). The world's storehouse of knowledge is readily available for those who read, and much of this information is not usually attained from other media (Comstock & Paik, 1991; Huston, Watkins, & Kunkel, 1989; Iyengar & Kinder, 1987; Postman, 1985; Zill & Winglee, 1990). Further, if we consider vocabulary to be one of the primary tools of verbal intelligence (Olson, 1986a), then we have another mechanism by which print exposure may influence cognition because reading appears to be a uniquely efficacious way of acquiring vocabulary (Hayes 1988, Hayes & Ahrens, 1988; Nagy & Anderson, 1984; Nagy & Herman, 1987).

In a study of forth-, fifth-, and sixth-grade children (Cunningham & Stanovich, 1991), we examined whether print exposure accounts for differences in vocabulary development once controls for both general and specific (i.e., vocabulary relevant) abilities are invoked. The analyses displayed in Table VII illustrate some of the outcomes of this study. Three different vocabulary measures were employed as dependent variables: a word checklist measure of the written vocabulary modeled on the work of Anderson and Freebody (1983; see also White et al., 1989; Zimmerman, Broder, Shaughnessy, & Underwood, 1977), a verbal fluency measure where the children had to output as many words as they could that fit into a particular category (e.g., things that are red, see Sincoff & Sternberg, 1987), a group-administered version of the Peabody Picture Vocabulary Test (PPVT). Age was entered first into the regression equation, followed by scores on the Raven Progressive Matrices as a control for general intelligence.

As a second ability control more closely linked to vocabulary acquisition mechanisms, we entered phonological coding ability into the equation. A variable such as phonological coding skill might mediate a relationship between print exposure and a variable like vocabulary size in numerous ways. High levels of decoding skill—certainly a contributor to greater print exposure—might provide relatively complete verbal contexts for the induction of word meanings during reading. Decoding skill might also indirectly reflect differences in short-term phonological storage that are related to vocabulary learning, particularly in the preschool years (Gathercole & Baddeley, 1989). Thus, print exposure and vocabulary might be spuriously linked via their connection with decoding ability: Good decoders read a lot and have the best context available for inferring new words.

TABLE VII

Unique Print Exposure Variance after Age, Raven, and Phonological Coding Were Partialed Out[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| | | *Word Checklist* | | |
| 1. Age | .103 | .011 | .011 | 1.41 |
| 2. Raven | .457 | .209 | .198 | 32.57** |
| 3. Phonological coding | .610 | .372 | .163 | 33.49** |
| 4. TRT | .683 | .466 | .094 | 22.52** |
| | | *Verbal Fluency* | | |
| 1. Age | .043 | .002 | .002 | 0.24 |
| 2. Raven | .231 | .053 | .051 | 6.89** |
| 3. Phonological coding | .477 | .228 | .175 | 28.47** |
| 4. TRT | .582 | .339 | .111 | 21.02** |
| | | *PPVT* | | |
| 1. Age | .230 | .053 | .053 | 7.29** |
| 2. Raven | .393 | .154 | .101 | 15.60** |
| 3. Phonological coding | .403 | .162 | .008 | 1.21 |
| 4. TRT | .516 | .266 | .104 | 18.19** |
| | | *Spelling* | | |
| 1. Age | .179 | .032 | .032 | 4.31* |
| 2. Raven | .414 | .172 | .140 | 21.95** |
| 3. Phonological coding | .656 | .430 | .258 | 58.51** |
| 4. TRT | .713 | .509 | .079 | 20.42** |
| | | *General Information* | | |
| 1. Age | .224 | .050 | .050 | 6.84** |
| 2. Raven | .362 | .131 | .081 | 12.05** |
| 3. Phonological coding | .410 | .168 | .037 | 5.68* |
| 4. TRT | .492 | .242 | .074 | 12.37** |

[a] The italic spanner headings identify the dependent variables in the regression analyses.
*$p < .05$. **$p < .01$.

This spurious linkage is controlled by entering phonological coding into the regression equation prior to the TRT. If print exposure were only an incidental correlate of vocabulary because of its linkage with phonological coding skill, then the TRT would not serve as a unique predictor of vocabulary once phonological coding was partialed out.

The results of the first three analyses displayed in Table VII indicate that for each of the vocabulary measures, the TRT accounted for significant variance

after the variance attributable to performance on the Raven and the phonological coding measure had been removed. The last two regressions indicate that this was also true for two additional criterion variables in the study: spelling ability and performance on the general information subtest of the WISC.

Similar relationships involving print exposure were found in a study of adult subjects (Stanovich & Cunningham, 1992). The first set of analyses, presented in Table VIII, partialed out general ability as measured by two nonverbal tasks before entering the print exposure measures as predictors. The dependent variables were a variety of indicators of verbal intelligence, including two vocabulary measures (the Nelson–Denny vocabulary subtest and the PPVT), a reading comprehension measure (Nelson–Denny), a measure of history and literature knowledge taken from the National Assessment of Educational Progress, a cultural literacy test, a composite measure of spelling performance, and a verbal fluency measure. The top half of Table VIII contains the cumulative *R*s resulting from forcing first figural analogies performance and then Raven matrices performance into the equation; two measures of print exposure (the ART and the MRT) were entered at the third step. The bottom half of the table shows the $R^2$ change values at each step of the analysis and whether the $R^2$ change values were significant at each step.

The results indicated that after performance on the figural analogies and Raven tasks was partialed out, the print exposure measures accounted for additional

TABLE VIII

Unique Print Exposure Variance after Nonverbal Abilities Are Partialed Out

| Step and variable | Dependent variables[a] | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | | | *Cumulative R* | | | |
| 1. Figural analogies | .316 | .278 | .280 | .270 | .332 | .238 | .205 |
| 2. Raven | .488 | .405 | .369 | .363 | .446 | .362 | .243 |
| 3. ART | .675 | .677 | .620 | .784 | .609 | .549 | .418 |
| 3. MRT | .628 | .599 | .564 | .720 | .548 | .438 | .325 |
| | | | | *$R^2$ Change* | | | |
| 1. Figural analogies | .100** | .077** | .079** | .073** | .110** | .057** | .042** |
| 2. Raven | .138** | .087** | .057** | .059** | .089** | .074** | .017* |
| 3. ART | .218** | .294** | .248** | .482** | .171** | .170** | .116** |
| 3. MRT | .156** | .194** | .182** | .386** | .102** | .061** | .047** |

[a] 1 = Nelson–Denny Vocabulary; 2 = PPVT; 3 = history and literature (NAEP); 4 = cultural literacy recognition; 5 = Nelson–Denny Comprehension; 6 = spelling composite; 7 = verbal fluency.

*$p < .05$. **$p < .001$.

variance in every dependent variable in the study. In many cases the unique variance explained was sizable. These analyses indicate that two indicators of print exposure can explain variance in verbal tasks not accounted for by general ability.

The next set of analyses provides a much more stringent test of the ability of the print exposure indicators to account for unique variance. The analyses in Table IX partial from the dependent variables reading comprehension ability in addition to the nonverbal ability measures. Performance on the Nelson–Denny reading comprehension subtest is entered in these hierarchical regressions subsequent to the two nonverbal ability tasks but prior to the measures of print exposure. Structuring the analyses in this way is not meant to imply that print exposure is not a determinant of reading comprehension ability. Indeed, there are strong grounds for believing that exposure to print *does* facilitate growth in comprehension ability (Anderson et al., 1988; Hayes, 1988; Juel, 1988; Stanovich, 1986). Thus, these analyses have allowed the Nelson–Denny comprehension measure to steal some of the variance that rightly belongs to the print exposure measures. The reason for structuring the analyses in this conservative manner was to ensure a stringent test of whether the print exposure measures

TABLE IX

Unique Print Exposure Variance after Nonverbal Abilities and Reading Comprehension Ability Are Partialed Out

| Step and variable | Dependent variables[a] | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | | | *Cumulative R* | | | |
| 1. Figural analogies | .316 | .278 | .280 | .270 | .238 | .205 |
| 2. Raven | .488 | .405 | .369 | .363 | .362 | .243 |
| 3. Nelson–Denny Comprehension | .684 | .541 | .599 | .600 | .582 | .323 |
| 4. ART | .738 | .688 | .677 | .803 | .625 | .423 |
| 4. MRT | .725 | .636 | .660 | .770 | .589 | .356 |
| | | | *$R^2$ Change* | | | |
| 1. Figural analogies | .100** | .077** | .079** | .073** | .057** | .042** |
| 2. Raven | .138** | .087** | .057** | .059** | .074** | .017* |
| 3. Nelson–Denny Comprehension | .230** | .129** | .222** | .227** | .208** | .045** |
| 4. ART | .076** | .180** | .100** | .286** | .052** | .075** |
| 4. MRT | .058** | .112** | .077** | .234** | .008 | .023* |

[a] 1 = Nelson–Denny Vocabulary; 2 = PPVT; 3 = history and literature (NAEP); 4 = cultural literacy recognition; 5 = spelling composite; 6 = verbal fluency.

*$p < .05$. **$p < .001$.

could predict performance on the criterion variables after possibly spurious relationships with general ability had been controlled.

The results illustrated in Table IX indicate that the ART was able to account for additional variance in all of the variables even after reading comprehension ability had been partialed out along with nonverbal ability. The MRT accounted for unique variance in four of five cases (the exception being spelling performance).

Overall, the analyses presented in Tables VIII and IX provide strong evidence that print exposure, independent of comprehension skill and other general abilities, is linked to vocabulary, verbal ability, and general knowledge. The print measures accounted for unique variance in vocabulary and general knowledge even after removal of variance explained by reading comprehension ability—an excellent measure of general verbal ability (Sternberg, 1987; Thorndike, 1973–1974).

The sample size in this study ($N = 300$) was large enough so that we were able to examine the consequences of a mismatch between general cognitive ability and print exposure. Although never losing sight of the correlational nature of the data, we may ask, for example, whether print exposure can compensate for modest levels of general cognitive abilities, at least in a statistical sense. The results of some relevant comparisons are presented in Table X. Two groups that were mismatched on print exposure and nonverbal cognitive ability were formed in the following manner. The sample was classified according to a median split of performance on the Raven matrices and on a composite print exposure score. The

TABLE X

Differences between Subjects with High Ability but Low in Print Exposure and Subjects with Low Ability but High in Print Exposure

| Variable | LoPrint/HiAbility[a] | HiPrint/LoAbility[b] | $t(116)$ |
|---|---|---|---|
| Raven matrices | 12.7 | 7.9 | −13.91** |
| Figural analogies | 13.1 | 12.7 | −0.93 |
| Author Recognition Test | .164 | .352 | 9.30** |
| Magazine Recognition Test | .433 | .605 | 9.00** |
| Nelson–Denny Vocabulary | 14.3 | 15.5 | 1.90 |
| Peabody Picture Vocabulary Test | 10.1 | 12.5 | 4.11** |
| History and literature (NAEP) | 12.1 | 13.9 | 3.24* |
| Cultural literacy recognition | .367 | .517 | 6.57** |
| Nelson–Denny Comprehension | 22.5 | 23.3 | 1.34 |
| Spelling composite | −.16 | .27 | 2.67* |
| Verbal fluency | 30.6 | 32.8 | 1.66 |

[a] $N = 56$.
[b] $N = 62$.
*$p < .05$. **$p < .001$.

resulting 2 × 2 matrix revealed 118 subjects who were discrepant: 56 subjects who were low in print exposure but high on the Raven (LoPrint/HiAbility) and 62 subjects who were high in print exposure but low on the Raven (HiPrint/LoAbility). These two groups were then compared on all of the variables in the study. Of course, significant differences were obtained on the variables that had defined the groups: the Raven, ART, and MRT. More interesting, however, is the fact that the HiPrint/LoAbility group was superior on all of the criterion variables in the study, significantly so in four cases.

Table XI displays an analysis of an even more unusual mismatch: that between print exposure and reading comprehension ability itself. Although it is assuredly the case that better readers read more, the correlation between the ability to read and the exercise of that ability is less than perfect. Some individuals read avidly despite modest skills, and others fail to exercise well-developed abilities. What are the cognitive correlates of a mismatch between abilities and the exercise of those abilities? To investigate this issue, the sample was classified according to a median split of performance on the Nelson–Denny comprehension subtest and a composite print exposure variable. The resulting 2 × 2 matrix revealed 82 subjects who were discrepant: 38 subjects who were low in print exposure but high in comprehension (LoPrint/HiComp) and 44 subjects who were high in print exposure but low in comprehension (HiPrint/LoComp). These two groups were then compared on all the variables in the study. Of course, significant differences were obtained on the variables that had defined the groups: the Nelson–Denny comprehension, ART, and MRT. The two groups were not differ-

TABLE XI

Differences between Subjects High in Comprehension Ability but Low in Print Exposure and Subjects Low in Comprehension Ability but High in Print Exposure

| Variable | LoPrint/HiComp[a] | HiPrint/LoComp[b] | $t(80)$ |
|---|---|---|---|
| Nelson–Denny Comprehension | 25.3 | 20.9 | −11.47** |
| Author Recognition Test | .186 | .310 | 5.27** |
| Magazine Recognition Test | .444 | .630 | 9.73** |
| Raven matrices | 10.7 | 9.0 | −2.44* |
| Figural analogies | 13.1 | 12.9 | −0.30 |
| Nelson–Denny Vocabulary | 15.1 | 14.4 | −0.94 |
| Peabody Picture Vocabulary Test | 10.6 | 12.1 | 2.06* |
| History and literature (NAEP) | 12.7 | 13.4 | 0.99 |
| Cultural literacy recognition | .396 | .483 | 3.86** |
| Spelling composite | .16 | −.05 | −1.12 |
| Verbal fluency | 31.6 | 32.0 | 0.30 |

[a] $N = 38$.
[b] $N = 44$.
*$p < .05$. **$p < .001$.

ent on the figural analogies measure; a significant difference favoring the LoPrint/HiComp group on the Raven matrices was obtained. However, despite comprehension differences favoring the LoPrint/HiComp group, as well as nonverbal cognitive abilities favoring this group (the Raven), LoPrint/HiComp individuals were not superior on any of the other variables. In fact, on two measures (the PPVT and cultural literacy test) the HiPrint/LoComp group performed significantly better. Print exposure appears to compensate for modest levels of general cognitive abilities, at least in a statistical sense. Although inferences from these correlational analysis must be tentative, the results do suggest that low ability need not necessarily hamper the development of vocabulary and verbal knowledge as long as the individual is exposed to a lot of print.

These data and those presented in our studies of children (see Tables IV and VII) refute the argument that experiential factors are not implicated or are of secondary importance in explaining performance on vocabulary measures. For example, Sternberg (1985) has argued that "Simply reading a lot does not guarantee a high vocabulary. What seems to be critical is not sheer amount of experience but rather what one has been able to learn from and do with that experience. According to this view, then, individual differences in knowledge acquisition have priority over individual differences in actual knowledge" (p. 307). Jensen (1980) has argued the point even more strongly, stating:

> Children of high intelligence acquire vocabulary at a faster rate than children of low intelligence, and as adults they have a much larger than average vocabulary, not primarily because they have spent more time in study or have been more exposed to words, but because they are capable of educing more meaning from single encounters with words. . . . The vocabulary test does not discriminate simply between those persons who have and those who have not been exposed to the words in context. . . . The crucial variable in vocabulary size is not exposure per se, but conceptual need and inference of meaning from context. (pp. 146–147)

The analyses reported here would seem to refute this argument if one accepts that the variables entered prior to print exposure in the analyses in Tables IV, VII, and IX are reasonable measures of general cognitive ability.

The data can be partitioned in additional ways that are informative on this issue. For example, reading comprehension is a general ability measure strongly related to skill in inducing word meanings. Sternberg (1987) so argues: "Whereas vocabulary is an indirect measure of ability to learn word meanings in context, reading comprehension is a fairly direct measure of ability to learn concepts in context. . . . The major difference would then be that reading comprehension tests measure present ability to learn from context, whereas vocabulary tests measure past ability" (p. 90). In Table XII are displayed the results of a communality analysis (Kerlinger & Pedhazur, 1973) of the overlap in variance among reading comprehension, print exposure, and vocabulary. Reading comprehension and print exposure overlap considerably in their variance shared with vocabulary measures (.227 and .295 for the PPVT and Nelson–Denny vocabulary, respec-

TABLE XII

Commonality Analyses

| | Unique | Common |
|---|---|---|
| *Vocabulary, Comprehension, and Print Exposure* | | |
| Dependent Variable: PPVT | | |
| Nelson–Denny Comprehension | .024 | .227 |
| Composite print exposure[a] | .220 | .227 |
| Dependent Variable: Nelson–Denny Vocabulary | | |
| Nelson–Denny Comprehension | .121 | .295 |
| Composite print exposure | .112 | .295 |
| *Vocabulary, General Ability, and Print Exposure* | | |
| Dependent Variable: PPVT | | |
| General ability[b] | .036 | .257 |
| Composite print exposure | .190 | .257 |
| Dependent Variable: Nelson–Denny Vocabulary | | |
| General ability | .149 | .319 |
| Composite print exposure | .088 | .319 |

[a] The composite print exposure variable is the average of the standard scores on the ART and MRT.

[b] General ability variance measured by the additive combination of variance on Raven matrices, figural analogies, and Nelson–Denny comprehension performance. The composite print exposure variable is the average of the standard scores on the ART and MRT.

tively); however, the print exposure measure explains as much unique variance as does reading comprehension on the Nelson–Denny vocabulary measure and explains considerably more unique variance than reading comprehension when the vocabulary measure is the PPVT. The bottom half of Table XII indicates that even when reading comprehension ability is amalgamated with the Raven and figural analogies tasks into a set of variables indexing general ability, the same relationships obtain. Print exposure remains separable from general ability and has as much unique predictive power as the ability composite.

### C. PRINT EXPOSURE AND INCIDENTAL VERSUS INTENTIONAL LEARNING

In a further study of college students (West & Stanovich, 1991), we attempted to sample knowledge domains that varied on the dimension of whether they reflected conscious, intentional learning of material in formal educational settings or whether they in part implicate the acquisition of information incidentally and informally in nonschool settings. This study also included SAT test scores as stringent control for spurious relationships involving general ability.

Our two key measures of formal, school learning were the students' college

grade point average and a content test on material from the subjects' major field. Our other two knowledge measures—a vocabulary measure and a cultural literacy test—are amalgamations of information acquired from formal schooling and from incidental learning in nonschool settings. Because the ART and MRT were designed to measure free-reading habits and not in-school study diligence, they would be expected to explain more unique variance on measures of knowledge acquired outside a formal school setting. We partialed out general ability as measured by SAT total scores (Table XIII) before entering the print exposure measures as predictors of the student's grade point average (GPA), the Area Concentration Achievement Test in Psychology (ACAT, Austin Peay State University, 1990), performance on a vocabulary checklist, and a cultural literacy test. The top half of Table XIII contains the cumulative *R*s; the bottom half shows the $R^2$ change values at each step and whether the $R^2$ change values were significant at each step.

The results for the four dependent variables diverged considerably. Neither of the print exposure measures predicted GPA or ACAT performance when entered after SAT performance, whereas each of the exposure measures accounted for significant additional variance on the vocabulary measure and on the cultural literacy test. This result is predictable if GPA and ACAT performance are assumed to reflect the intentional learning of school material. In contrast, the vocabulary and cultural literacy measures reflect the amalgamation of information acquired from formal schooling and from incidental learning in nonschool

TABLE XIII

Hierarchical Regression Analyses

| Step and variable | Dependent variables[a] | | | |
|---|---|---|---|---|
| | GPA | ACAT | Vocab | CLT |
| | | *Cumulative R* | | |
| 1. SAT | .343 | .306 | .510 | .381 |
| 2. ART | .345 | .306 | .564 | .570 |
| 2. MRT | .360 | .348 | .583 | .558 |
| | | $R^2$ *change* | | |
| 1. SAT | .118* | .093* | .260* | .145* |
| 2. ART | .001 | .000 | .058* | .180* |
| 2. MRT | .011 | .028 | .080* | .167* |

[a] GPA = grade point average, ACAT = Area Concentration Achievement Test in Psychology, Vocab = vocabulary checklist, CLT = cultural literacy test.

*$p < .01$.

settings. The obtained results, then, occur because the exposure measures in part reflect nonschool information acquisition.

## D. THE RECOGNITION CHECKLISTS AND READING IN THE REAL WORLD

In another study (West, Stanovich, & Mitchell, in press) we attempted to validate the checklist print exposure measures by seeing whether they were associated with individual differences in reading observed in a nonlaboratory setting where reading occurs. The setting chosen for our study was an airport passenger waiting lounge. This is a setting where reading occurs via the free choice of the subject. If individual differences in free reading in a setting such as this can be related to performance on the recognition checklist tasks, this would strongly bolster the construct validity of the checklist measures as indicators of individual differences in print exposure.

The study involved unobtrusive observations of individuals in a waiting lounge at National Airport in Washington, DC. Individuals sitting by themselves were the potential subjects. Such individuals were selected and monitored unobtrusively for 10 consecutive minutes. If they were not reading at the beginning of the observation period and continued sitting by themselves without reading or having reading matter in sight for the entire 10-minute period, they were classified as nonreaders. If they were reading at the beginning of the observation period and continued reading for the entire 10-minute period they were classified as readers. Individuals whose behavior did not fall into one of these categories did not enter the sample. Subsequent to the observation, the individual was approached by the experimenter, was asked for consent to participate in the study and to fill out several experimental measures, and then was debriefed. Slightly less than 10% of the potential subjects refused to participate.

Table XIV displays the results of a comparison of the readers ($N = 111$) and nonreaders ($N = 106$) on a few of the measures. The groups were significantly different on the ART, the MRT, and a newspaper recognition test; however, they were not different on measures of exposure to television and film. This pattern of differences provides evidence of ecological validity for the recognition checklist measures. They were reliably linked to direct observations of the behavior of interest (free reading) in a situation where investigators do not intrude on the process. The trends were not due to just a few items on the tasks. For example, airport readers displayed significantly greater recognition of over 80% of the items on the ART.

Importantly, the readers were also superior on measures of vocabulary and general knowledge (a cultural literacy recognition test); however, as the last two rows in Table XIV show, the readers were also older and had more education.

TABLE XIV

Differences between Readers and Nonreaders[a]

| Variable | Nonreaders | Readers | *t* value |
|---|---|---|---|
| Author Recognition Test | .401 | .635 | 7.75* |
| Magazine Recognition Test | .598 | .751 | 5.21* |
| Newspaper Recognition Test | .370 | .529 | 6.12* |
| Television Recognition Test | .426 | .468 | 1.87 |
| Film Recognition Test | .292 | .320 | 1.10 |
| Vocabulary checklist | .516 | .731 | 7.57* |
| Cultural literacy recognition | .600 | .770 | 7.00* |
| Age | 35.3 | 41.4 | 3.28* |
| Education | 15.2 | 16.5 | 4.25* |

[a] $df = 211$ for the vocabulary checklist, 213 for the MRT, 214 for film recognition, and 215 for all other variables.

*$p < .01$.

Although the correlations of reading with age and education probably present real relationships in the population and should not be interpreted as confounds, we nevertheless carried out analyses designed to examine whether airport reading was related to the print exposure measures after the effects of age and education had been controlled. Table XV presents the results of three hierarchical regression analyses in which age and education were entered as predictors of airport reading (scored as a 0/1 variable) prior to each of the three recognition checklist measures of print exposure. Age was significantly related to airport reading. Education was a significant predictor after the variance explained by age had been removed; however, each of the three measures of print exposure was significantly linked to airport reading even after variance explained by both age and education was removed.

Although age and education were not mediating the relationship between airport reading and the other print exposure measures, it is still possible that age and/or education might have resulted in a spurious linkage between airport reading and performance on the vocabulary and literacy measures. The results of the two hierarchical regressions presented in Table XVI address this possibility. In these regressions, age and education were entered prior to airport reading (scored dichotomously) as predictors of vocabulary and general knowledge. In both analyses, airport reading remained a significant predictor after age and education had been controlled. The variance explained here was less than that in some of our earlier analyses, but of course in those studies, we had spent prodigious amounts of time and effort constructing the measures of print exposure. Here we have a "10-minute airport test" that serves as a vocabulary predictor independent of educational level!

TABLE XV

Hierarchical Regressions Predicting Airport Reading[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| 1. Age | .218 | .048 | .048 | 10.75* |
| 2. Education | .329 | .108 | .060 | 14.57* |
| 3. ART | .503 | .253 | .145 | 41.17* |
| 1. Age | .227 | .051 | .051 | 11.55* |
| 2. Education | .342 | .117 | .066 | 15.65* |
| 3. MRT | .426 | .181 | .064 | 16.69* |
| 1. Age | .218 | .048 | .048 | 10.75* |
| 2. Education | .329 | .108 | .060 | 14.57* |
| 3. Newspaper recognition | .444 | .198 | .090 | 23.69* |

[a] $N = 217$ in the first analysis, 215 in the second, and 217 in the third.
*$p < .01$.

Not surprisingly, given our earlier results, the checklist measures of print exposure were also significant predictors of vocabulary after age and education were partialed out. In fact, the analyses displayed in Table XVII illustrate that the print exposure measures remain unique predictors even after another control is invoked: entering television recognition as the third variable in the equation. This

TABLE XVI

Airport Reading as a Predictor of Vocabulary and Cultural Literacy[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| | *Vocabulary Checklist* | | | |
| 1. Age | .257 | .066 | .066 | 14.91* |
| 2. Education | .562 | .315 | .249 | 76.52* |
| 3. Airport reading | .638 | .408 | .093 | 32.50* |
| | *Cultural Literacy Test* | | | |
| 1. Age | .211 | .045 | .045 | 10.04* |
| 2. Education | .495 | .245 | .200 | 56.65* |
| 3. Airport reading | .574 | .329 | .084 | 34.81* |

[a] $N = 213$ in the first analysis and 217 in the second.
*$p < .01$.

TABLE XVII

Hierarchical Regressions Predicting Vocabulary Scores[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| 1. Age | .257 | .066 | .066 | 14.91* |
| 2. Education | .562 | .315 | .249 | 76.52* |
| 3. TV recognition | .607 | .369 | .054 | 17.57* |
| 4. ART | .729 | .531 | .162 | 72.12* |
| 1. Age | .253 | .064 | .064 | 14.27* |
| 2. Education | .555 | .308 | .244 | 73.34* |
| 3. TV recognition | .604 | .365 | .057 | 18.42* |
| 4. MRT | .659 | .434 | .069 | 25.44* |
| 1. Age | .257 | .066 | .066 | 14.91* |
| 2. Education | .562 | .315 | .249 | 76.52* |
| 3. TV recognition | .607 | .369 | .054 | 17.57* |
| 4. Newspaper recognition | .663 | .439 | .070 | 26.25* |

[a] $N = 213$ in the first analysis, 211 in the second, and 213 in the third.
*$p < .01$.

variable not only controls for any variance in vocabulary knowledge that is specifically linked to television, but it also serves as a control for method variance because it has exactly the same checklist logic and response requirements as the print exposure measures. Table XVIII illustrates that the print exposure measures also remain significant predictors of performance on the cultural literacy measure.

This study successfully demonstrated the ability of the recognition checklist measures to predict reading behavior in a real-life setting, one where the target behavior is not influenced by the presence of an experimenter. The construct validity of the tasks has now been bolstered by demonstrating linkages with other measures of print exposure (e.g., the diary method, see Allen et al., 1992) and by demonstrating their ability to predict behavior in a natural environment.

Print exposure, whether measured by the 10-minute airport probe or by the recognition checklist measures, was significantly linked to vocabulary and cultural knowledge even after controls for age and education were invoked. This finding converges with the previously reported studies that have indicated that print exposure can predict a variety of behavioral outcomes even when some rather stringent controls for general cognitive ability and background characteristics are employed.

TABLE XVIII

Hierarchical Regressions Predicting Cultural Literacy Scores[a]

| Step and variable | Statistic | | | |
|---|---|---|---|---|
| | $R$ | $R^2$ | $R^2$ change | $F$ to enter |
| 1. Age | .211 | .045 | .045 | 10.04* |
| 2. Education | .495 | .245 | .200 | 56.65* |
| 3. TV recognition | .654 | .427 | .182 | 67.96* |
| 4. ART | .787 | .620 | .193 | 107.19* |
| 1. Age | .208 | .043 | .043 | 9.67* |
| 2. Education | .486 | .236 | .193 | 53.51* |
| 3. TV recognition | .653 | .427 | .191 | 70.03* |
| 4. MRT | .736 | .542 | .115 | 52.88* |
| 1. Age | .211 | .045 | .045 | 10.04* |
| 2. Education | .495 | .245 | .200 | 56.65* |
| 3. TV recognition | .654 | .427 | .182 | 67.96* |
| 4. Newspaper recognition | .712 | .507 | .080 | 34.28* |

[a] $N = 217$ in the first analysis, 215 in the second, and 217 in the third.
*$p < .01$.

## VIII. Summary and Conclusions

The studies reported here represent the first steps in the development of a new research paradigm for studying the unique cognitive correlates of literacy. Reading experience exhibits enough isolable variance within a generally literate society to be reliably linked with cognitive differences. Research on such links is therefore facilitated because the consequences of engaging in literacy activities can be studied without necessarily obtaining totally illiterate samples or setting up cross-cultural comparisons. Issues that are at least analogous issues to those raised in cross-cultural research can be studied within literate societies with a paradigm such as this, and therefore the speed with which we can answer questions about the cognitive consequences of literacy may be greatly increased because more studies can be carried out, larger samples can be studied, and the range of the cognitive domains tapped can be widened.

Research in this area appears to have been stifled because of the widespread acceptance of the most extreme interpretations of the outcome of Scribner and Cole's (1981) investigation—interpretations that have slowly diffused throughout the literature without being accompanied by any new data. These conclusions are fueled by a powerful social critique that advances the argument that the

positive cultural and economic effects of literacy have been overstated—indeed, that literacy is, if anything, a repressive force (Auerbach, 1992; Street, 1984, 1988; Stuckey, 1991). Educational theorists such as Frank Smith accused the educational establishment of "overselling" literacy and have argued that "Literacy doesn't generate finer feelings or higher values. It doesn't even make anyone smarter" (1989, p. 354).

The data reported herein appear to indicate that these theorists could well be wrong in this conclusion. If "smarter" means having a larger vocabulary and more world knowledge in addition to the abstract reasoning skills encompassed within the concept of intelligence, as it does in most laymen's definitions of intelligence (Stanovich, 1989; Sternberg, 1990), then reading may *well* make people smarter. Certainly our data demonstrate time and again that print exposure is associated with vocabulary, general knowledge, and verbal skills even after controlling for abstract reasoning abilities (as measured by such indicators as the Raven). Although nothing can turn our correlational data into true experimental findings, the converging patterns of relationships—most importantly the indication that reading habits predict growth in verbal abilities in longitudinal investigations (see Tables V and VI; Anderson et al., 1988; Juel, 1988)—certainly imply a role for reading experience in a comprehensive theory of cognitive growth (its role is at least as well supported as many other mechanisms that have attained popularity in developmental psychology).

Thus, investigators who attempt to supplement purely genetic accounts of differences in mental ability by speculating about variables in children's ecologies that could account for cognitive change (e.g., Ceci, 1990) might well find print exposure worth investigating, because the variables they choose must have the requisite potency to perform their theoretical roles. An important class of such variables would be those that have long-term effects because of their repetitive and/or cumulative action. Schooling is obviously one such variable (Cahan & Cohen, 1989; Ceci, 1990, 1991; Morrison, 1987); however, print exposure is another variable that cumulates over time into enormous individual differences. For example, Anderson et al. (1988) have found hundredfold differences in word exposure among fifth-grade (10- to 11-year-old) children and order-of-magnitude differences in opportunities to learn vocabulary words (see also Hayes & Ahrens, 1988). From the time of at least the fifth grade, an avid reader is seeing literally millions of words a year (Anderson et al., 1988). Thus, whatever cognitive processes are engaged over word or word-group units (phonological coding, semantic activation, parsing, induction of new vocabulary items) are being exercised hundreds of times a day. This amount of cognitive muscle flexing might be expected to have some specific effects. Reading volume is thus an explanatory variable that should be more routinely considered when attempting to predict individual cognitive outcomes and group trends. For example, print exposure might be a useful explanatory variable that can be called on

when trying to explain group trends such as declining verbal SAT scores (Wirtz, 1977), historical changes in intelligence test performance (Flynn, 1987), or differential changes in fluid versus crystallized intelligence with aging.

The results reported here do not, of course, reveal anything about the *causes* of differences in exposure to print. Certainly, environmental differences (cultural opportunities, parental modeling, quality of schooling) may be a contributing factor (Anderson et al., 1988). But personality dispositions toward literacy activities may also play a role, and the environmental and/or genetic determinants of such behavioral propensities are completely unknown (but see Plomin, Corley, DeFries, & Fulker, 1990). We must be careful to avoid the "sociologist's fallacy" of failing to recognize that a seemingly environmental variable like print exposure could, via the influence of the parent-constructed home literacy environment, carry genetic variance (Plomin & Bergeman, 1991). Nevertheless, Olson (1991) analyzed the heritability of the deficit in performance on the print exposure checklist measures shown by dyslexic twins in the Colorado Reading Project and found that the hypothesis of zero heritability could not be rejected.

What are the mechanisms by which print exposure comes to be an independent predictor of variance in the criterion variables studied in this investigation? Several mechanisms are possible. First, the distributions of language structures that people are exposed to in print are different from those encountered in speech. Evidence for this conjecture is most strong in the lexical domain. Work by Hayes (1988; Hayes & Ahrens, 1988; see also Akinnaso, 1982; Biber, 1986; Chafe & Danielewicz, 1987; Corson, 1985) has indicated that moderate- to low-frequency words—precisely those words that differentiate individuals with high and low vocabulary sizes—appear much more often in common reading matter than in common speech. These relative differences in the statistical distributions of words in print and in oral language have direct implications for vocabulary development.

Most theorists agree that a substantial proportion of vocabulary growth during childhood occurs indirectly through language exposure (Miller & Gildea, 1987; Nagy & Anderson, 1984; Nagy, Herman, & Anderson, 1985; Sternberg, 1985, 1987). Furthermore, many researchers are convinced that exposure to print is a more potent source of vocabulary growth than is exposure to oral language (Hayes, 1988; Hayes & Ahrens, 1988; Krashen, 1989; Nagy & Anderson, 1984; Nagy & Herman, 1987; Stanovich, 1986). If most of one's vocabulary is acquired outside formal teaching, then the only opportunities to acquire new words occur when an individual is exposed to a word in written or oral language that is outside the current vocabulary. That such exposure will happen vastly more often while reading than while talking or watching television is illustrated in research by Hayes and Ahrens (1988). They studied how many rare words per 1000 are contained in various categories of language. A rare word was defined as one with a rank lower than 10,000 in the Carroll, Davies, and Richman (1971) count,

roughly, a word that is outside the vocabulary of forth to sixth graders (9- to 12-year-olds). For vocabulary growth to occur after the middle grades, children must be exposed to words that are rare by this definition. Hayes and Ahrens (1988) found that print provides many more such word-learning opportunities. Children's books contain 50% more rare words than do adult prime-time television and the conversation of college graduates. Popular magazines have roughly three times as many opportunities for new word learning as prime-time television and adult conversation. The data presented by Hayes and Ahrens (1988) indicate that conversation is not a substitute for reading.

To a lesser extent, a similar situation holds for other language systems, including syntax (Purcell-Gates, 1988). Although all syntactic constructions can be found in all types of language, more complex syntactic constructions are disproportionately found in text (Akinnaso, 1982; Biber, 1986; Redeker, 1984). Of course, complex syntactic constructions are also found disproportionately in types of speech that are textlike, such as judicial proceedings, planned speeches, and college lectures (Biber, 1986; Chafe & Danielewicz, 1987); nevertheless, the findings establish that the average person experiences these syntactic constructions disproportionately in print (Chafe & Danielewicz, 1987; Purcell-Gates, 1988; Redeker, 1984). In short, print exposure might be expected to contribute to skill in verbal domains because print is a source of exceptionally rich stimulation.

Another mechanism by which print exposure might lead to cognitive change is its role as a builder of the individual's knowledge base. In recent years, cognitive and developmental psychologists have strongly emphasized the importance of domain knowledge in determining information processing efficiency (Bjorklund, 1987; Ceci, 1990; Chi, Hutchinson, & Robin, 1989; Keil, 1984; Scribner, 1986). Yussen (1990) summarized the view as follows: "Much of what develops in children's memory is neither changes in basic capacity nor changes in strategies available to children but, instead, the richness of knowledge about a topic or about the concepts embedded in the material put to children to remember in various experimental tasks. This is called content" (p. 677). Print is a uniquely rich source of content. Personal experience provides only narrow knowledge of the world and is often misleading and unrepresentative (Baron, 1985, 1988; Dawes, 1988; Gilovich, 1991; Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980). The most commonly used electronic sources of information (television, radio) lack depth (Comstock & Paik, 1991; Hayes & Ahrens, 1988; Huston et al., 1989; Iyengar & Kinder, 1987; Zill & Winglee, 1990). Only print provides opportunities for acquiring broad and deep knowledge of the world. Research indicates that reading has higher correlations with world and cultural knowledge than does television viewing (Allen et al., 1992; West et al., 1992; West & Stanovich, 1991; Zill & Winglee, 1990).

Cognitive theories in which individual differences in basic processing capacities are viewed as at least partly determined by differences in knowledge bases

(e.g., Ceci, 1990) indirectly provide a mechanism through which print exposure influences cognitive efficiency. Print is simply a more distal factor that determines individual differences in knowledge bases, which in turn influence performance on a variety of basic information processing tasks (see Ceci, 1990). This link explains why some of the relations found in our studies between print exposure and criterion variables such as general knowledge and vocabulary should not be criticized or dismissed as representing "narrow" effects. If the theories of cognitive development in which domain knowledge is emphasized have some truth to them, then demonstrating effects on such knowledge structures is an important finding, because whatever causal power accrues to content knowledge in these theories also partially accrues to print exposure as a mechanism of cognitive change.

Finally, in any attempt to explain tendencies of early achievement disparities to increase with age (e.g., Jorm, Share, Maclean, & Matthews, 1984; Stanovich, 1986), print exposure is, again, a variable that may have some explanatory power. For example, Hayes and Grether (1983) studied the growth in reading comprehension and vocabulary of several thousand students in the New York City schools during the school year and during the summer. They found that the summer period, when the children were not in school, accounted for more of the gap between the high-achieving and low-achieving students than did the period when the children were actually in school. They concluded that "it now appears that non-school periods may have contributed a majority of the differentials in reading and word knowledge noted among the six sets of schools" (pp. 65–66). "In short, very little of the enormous difference in word knowledge performance . . . appears to be attributable to what goes on *in* school; most of it comes from what goes on *out* of school" (p. 64). Demonstrating that the pattern uncovered by Hayes and Grether can be explained in terms of specific variables in children's ecologies would be important, and print exposure might play a role in such an explanation.

The Hayes and Grether result may represent an instance of "Matthew effects" in literacy development: educational sequences in which early and efficient acquisition of reading skill yields faster rates of growth in reading achievement and other cognitive skills (see Stanovich, 1986; Walberg & Tsai, 1983). The term derives from the Gospel according to Matthew—"For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath" (XXV:29)—and refers to rich-get-richer and poor-get-poorer effects embedded in the sociodevelopmental context of schooling. For example, children who are already good comprehenders may tend to read more, thus spurring further increases in their reading comprehension abilities and increasing the achievement differences between them and their agemates who are not good comprehenders and not avid readers (Chall, Jacobs & Baldwin, 1990; Juel, 1988; Share & Silva, 1987; Share, McGee, & Silva, 1989;

Stanovich, 1986; van den Bos, 1989). Thus, free-reading choices may explain part of the puzzle—and the pressing social problem—of widening achievement disparities between the educational haves and have-nots.

## ACKNOWLEDGMENT

Research reported in this article was supported by a grant from the Social Sciences and Humanities Research Council of Canada.

## REFERENCES

Akinnaso, F. N. (1981). The consequences of literacy in pragmatic and theoretical perspectives. *Anthropology & Education Quarterly, 12,* 163–200.

Akinnaso, F. N. (1982). On the difference between spoken and written language. *Language and Speech, 25,* 97–125.

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84,* 489–503.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Huston (Ed.), *Advances in reading/language research* (Vol. 2, pp. 231–256). Greenwich, CT: JAI Press.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23,* 285–303.

Applebee, A. N., Langer, J. A., & Mullis, I. V. S. (1988). *Who reads best?* Princeton, NJ: Educational Testing Service.

Auerbach, E. (1992). Literacy and ideology. *Annual Review of Applied Linguistics, 12,* 71–85.

Austin Peay State University. (1990). *Area Concentration Achievement Test in Psychology.* Clarksville, TN: Project for Area Concentration Achievement Tests, Austin Peay State University.

Baron, J. (1985). *Rationality and intelligence.* Cambridge: Cambridge University Press.

Baron, J. (1988). *Thinking and deciding.* Cambridge: Cambridge University Press.

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62,* 384–414.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7,* 93–130.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science, 236,* 157–161.

Burt, C. D. B., & Kemp, S. (1991). Retrospective duration estimation of public events. *Memory & Cognition, 19,* 252–262.

Byrne, B. (1981). Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics, 2,* 201–212.

Cahan, S , & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development, 60,* 1239–1249.

Carp, F. M., & Carp, A. (1981). The validity, reliability and generalizability of diary data. *Experimental Aging Research, 7,* 281–296.

Carr, T. H., & Levy, B. A. (Eds.). (1990). *Reading and its development: Component skills approaches.* San Diego: Academic Press.

Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book.* Boston: Houghton Mifflin.

Ceci, S. J. (1990). *On intelligence . . . more or less: A bio-ecological treatise on intellectual development.* Englewood Cliffs, NJ: Prentice-Hall.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27,* 703–722.

Chafe, W., & Danielewicz, J. (1987). Properties of spoken and written language. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 83–113). San Diego: Academic Press.

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind.* Cambridge, MA: Harvard University Press.

Chen, C., & Stevenson, H. W. (1989). Homework: A cross-cultural examination. *Child Development, 60,* 551–561.

Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly, 35,* 27–62.

Chomsky, C. (1972). Stages in language development and reading exposure. *Harvard Educational Review, 42,* 1–33.

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.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement.* New York: Basic Books.

Comstock, G., & Paik, H. (1991). *Television and the American child.* San Diego: Academic Press.

Cooksey, R. W., & Freebody, P. (1987). Aspects of a computer-managed test of children's reading vocabulary: Reliability, validity, and characterization of knowledge. *Reading Psychology, 8,* 103–118.

Corson, D. (1985). *The lexical bar.* Oxford: Pergamon.

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.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83,* 264–274.

Cunningham, A. E., Stanovich, K. E., Wilson, M. R. (1990). Cognitive variation in adult students differing in reading ability. In T. Carr & B. A. Levy (Eds.), *Reading and development: Component skills approaches* (pp. 129–159). San Diego: Academic Press.

Curtis, M. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72,* 656–669.

Dawes, R. M. (1988). *Rational choice in an uncertain world.* San Diego: Harcourt Brace Jovanovich.

Ennis, P. H. (1965). *Adult book reading in the United States.* National Opinion Research Center Report No. 105. Chicago: University of Chicago.

Erickson, F. (1984). School literacy, reasoning, and civility: An anthropologist's perspective. *Review of Educational Research, 54,* 525–546.

Estes, T. H. (1971). A scale to measure attitudes toward reading. *Journal of Reading, 15,* 135–138.

Feldman, C. F. (1991). Oral metalanguage. In D. R. Olson & N. Torrance (Eds.), *Literacy and orality* (pp. 47–65). Cambridge: Cambridge University Press.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191.

Fuller, B., Edwards, J., & Gorman, K. (1987). Does rising literacy spark economic growth? Commercial expansion in Mexico. In D. A. Wagner (Ed.), *The future of literacy in a changing world* (pp. 319–340). Oxford: Pergamon.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality & Individual Differences, 7,* 385–400.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study, *Journal of Memory and Language, 28,* 200–213.

Gee, J. P. (1988). The legacies of literacy: From Plato to Freire through Harvey Graff. *Harvard Educational Review, 58,* 195–212.

Gellner, E. (1985). *Relativism and the social sciences.* Cambridge: Cambridge University Press.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences.* San Francisco: Freeman.

Gilovich, T. (1991). *How we know what isn't so.* New York: Free Press.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39,* 93–104.

Goody, J. (1977). *The domestication of the savage mind.* New York: Cambridge University Press.

Goody, J. (1980). Thought and writing. In E. Gellner (Ed.), *Soviet and Western anthropology* (pp. 119–133). London: Duckworth.

Goody, J. (1987). *The interface between the written and the oral.* Cambridge: Cambridge University Press.

Goody, J., & Watt, I. (1968). The consequences of literacy. In J. Goody (Ed.), *Literacy in traditional societies* (pp. 27–68). London: Cambridge University Press.

Graff, H. J. (1979). *The literacy myth.* New York: Academic Press.

Graff, H. J. (1986). The legacies of literacy: Continuities and contradictions in Western society and culture. In S. de Castell, A. Luke, & K. Egan (Eds.), *Literacy, society, and schooling* (pp. 61–86). Cambridge: Cambridge University Press.

Graff, H. J. (1987). *The labyrinths of literacy.* London: Falmer.

Greaney, V. (1980). Factors related to amount and time of leisure time reading. *Reading Research Quarterly, 15,* 337–357.

Greaney, V., & Hegarty, M. (1987). Correlates of leisure-time reading. *Journal of Research in Reading, 10,* 3–20.

Greenfield, P. (1972). Oral or written language: The consequences for cognitive development in Africa, the United States and England. *Language and Speech, 15,* 169–178.

Guthrie, J. T. (1981). Reading in New Zealand: Achievement and volume. *Reading Research Quarterly, 17,* 6–27.

Guthrie, J. T., & Greaney, V. (1991). Literacy acts. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 68–96). New York: Longman.

Guthrie, J. T., Schafer, W. D., & Hutchinson, S. R. (1991). Relations of document literacy and prose literacy to occupational and societal characteristics of young black and white adults. *Reading Research Quarterly, 26,* 30–48.

Guthrie, J. T., & Seifert, M. (1983). Profiles of reading activity in a community. *Journal of Reading, 26,* 498–508.

Havelock, E. A. (1963). *Preface to Plato.* Cambridge: Harvard University Press.

Havelock, E. A. (1980). The coming of literate communication to Western culture. *Journal of Communication, 30,* 90–98.

Hayes, D. P. (1988). Speaking and writing: Distinct patterns of word choice. *Journal of Memory and Language, 27,* 572–585.

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language, 15,* 395–410.

Hayes, D. P., & Grether, J. (1983). The school year and vacations: When do students learn? *Cornell Journal of Social Relations, 17*(1), 56–71.

Hollis, M., & Lukes, S. (Eds.) (1982). *Rationality and relativism.* Cambridge, MA: MIT Press.

Huston, A., Watkins, B. A., & Kunkel, D. (1989). Public policy and children's television. *American Psychologist, 44,* 424–433.

Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American opinion.* Chicago: University of Chicago Press.

Jensen, A. (1980). *Bias in mental testing.* New York: Free Press.

Jorm, A., Share, D., Maclean, R., & Matthews, R. (1984). Phonological recoding skills and learning to read: A longitudinal study. *Applied Psycholinguistics, 5,* 201–207.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447.

Just, M., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Boston: Allyn & Bacon.

Kaestle, C. F. (1991). *Literacy in the United States.* New Haven, CT: Yale University Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kamhi, A., & Catts, H. (1989). *Reading disabilities: A developmental language perspective.* Boston: College-Hill Press.

Keil, F. C. (1984). Mechanisms of cognitive development and the structure of knowledge. In R. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 81–99). New York: Freeman.

Kerlinger, F. N., & Pedhazur, E. J. (1973). *Multiple regression in behavioral research.* New York: Holt, Rinehart & Winston.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73,* 440–464.

Lewis, R., & Teale, W. H. (1980). Another look at secondary school students' attitudes toward reading. *Journal of Reading Behavior, 12,* 187–201.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Manning, M. (1988). *The standard periodical directory* 11th ed.). New York: Oxbridge Communications.

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher, 43,* 626–639.

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing, 4,* 142–151.

Miller, G. A., & Gildea, P. M. (1987). How children learn words. *Scientific American, 257*(3), 94–99.

Morrison, F. J. (1987, November). *The "5–7" shift revisited: A natural experiment.* Paper presented at the meeting of the Psychonomic Society, Seattle, WA.

Musgrove, F. (1982). *Education and anthropology.* New York: Wiley.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19,* 304–330.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–35). Hillsdale, NJ: Erlbaum.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20,* 233–253.

Nisbett, L., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Nystrand, M. (1987). The role of context in written communication. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 197–214). San Diego: Academic Press.

Oakhill, J., & Garnham, A. (1988). *Becoming a skilled reader.* Oxford: Basil Blackwell.

Olson, D. R. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review, 47,* 257–281.

Olson, D. R. (1986a). Intelligence and literacy: The relationships between intelligence and the technologies of representation and communication. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence* (pp. 338–360). Cambridge: Cambridge University Press.

Olson, D. R. (1986b). The cognitive consequences of literacy. *Canadian Psychology, 27,* 109–121.

Olson, D. R. (1988). Mind and media: The epistemic functions of literacy. *Journal of Communication, 38,* 27–36.

Olson, D. R., & Torrance, N. (Eds.) (1991). *Literacy and orality.* Cambridge, England: Cambridge University Press.

Olson, R. K. (1991, September). *Genetic etiologies of reading disability.* Paper presented at the NATO Advanced Study Institute on Differential Diagnosis and Treatments of Reading and Writing Disorders, Chateau de Bonas, France.

Ong, W. J. (1967). *The presence of the word.* Minneapolis: University of Minnesota Press.

Ong, W. J. (1982). *Orality and literacy.* London: Methuen.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language, 24,* 59–88.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influences on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373–427.

Plomin, R., Corley, R., DeFries, J. C., & Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. *Psychological Science, 1,* 371–377.

Plomin, R., DeFries, J. C., & McClearn, G. E. (1990). *Behavioral genetics: A primer* (2nd ed.). New York: Freeman.

Postman, N. (1985). *Amusing ourselves to death.* New York: Viking Penguin.

Purcell-Gates, V. (1988). Lexical and syntactic knowledge of written narrative held by well-read-to kindergartners and second graders. *Research in the Teaching of English, 22,* 128–157.

Ravitch, D., & Finn, C. E. (1987). *What do our 17-year-olds know?* New York: Harper & Row.

Redeker, G. (1984). On differences between spoken and written language. *Discourse Processes, 7,* 43–55.

Rice, G. E. (1986). The everyday activities of adults: Implications for prose recall, I. *Educational Gerontology, 12,* 173–186.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence* (pp. 13–30). Cambridge: Cambridge University Press.

Scribner, S., & Cole, M. (1978). Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review, 48,* 448–461.

Scribner, S., & Cole, M. (1981). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Share, D. L., McGee, R., & Silva, P. (1989). IQ and reading progress: A test of the capacity notion of IQ. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 97–100.

Share, D. L., & Silva, P. A. (1987). Language deficits and specific reading retardation: Cause or effect? *British Journal of Disorders of Communication, 22,* 219–226.

Sharon, A. T. (1973–1974). What do adults read? *Reading Research Quarterly, 9,* 148–169.

Shweder, R. A. (1991). *Thinking through cultures.* Cambridge, MA: Harvard University Press.

Siegel, H. (1988). *Educating reason.* New York: Routledge.

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., & Ryan, E. B. (1989). Subtypes of developmental dyslexia: The influence of definitional variables. *Reading and Writing: An Interdisciplinary Journal, 1,* 257–287.

Sincoff, J. B., & Sternberg, R. J. (1987). Two faces of verbal ability. *Intelligence, 11,* 263–276.

Smith, F. (1989). Overselling literacy. *Phi Delta Kappan, 70*(5), 353–359.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Sperber, D. (1985). Apparently irrational beliefs. In D. Sperber (Ed.), *On anthropological knowledge* (pp. 35–63). Cambridge: Cambridge University Press.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–407.

Stanovich, K. E. (1989). Has the learning disabilities field lost its intelligence? *Journal of Learning Disabilities, 22,* 487–492.

Stanovich, K. E., & Cunningham, A. E. (1991). Reading as constrained reasoning. In R. Sternberg & P. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 3–60). Hillsdale, NJ: Erlbaum.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20,* 51–68.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19,* 278–303.

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

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence.* Cambridge: Cambridge University Press.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89–105). Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* Cambridge: Cambridge University Press.

Stevenson, H. W., Stigler, J. W., Lee, S. Y., Lucker, G. W., Kitamura, S., & Hsu, C. C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development, 56,* 718–734.

Street, B. V. (1984). *Literacy in theory and practice.* Cambridge: Cambridge University Press.

Street, B. V. (1988). Literacy practices and literacy myths. In R. Saljo (Ed.), *The written word: Studies in literate thought and action* (pp. 59–72). Berlin: Springer-Verlag.

Stuckey, J. E. (1991). *The violence of literacy.* Portsmouth, NH: Boynton/Cook.

Tannen, D. (1982). The myth of orality and literacy. In W. Frawley (Ed.), *Linguistics and literacy* (pp. 37–50). New York: Plenum Press.

Taylor, B. M., Frye, B. J., & Maruyama, G. M. (1990). Time spent reading and reading growth. *American Educational Research Journal, 27,* 351–362.

Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science, 2,* 158–165.

Thorndike, R. L. (1973–1974). Reading as reasoning. *Reading Research Quarterly, 9,* 135–147.

Van den Bos, K. P. (1989). Relationship between cognitive development, decoding skill, and reading comprehension in learning disabled Dutch children. In P. Aaron & M. Joshi (Eds.), *Reading and writing disorders in different orthographic systems* (pp. 75–86). Dordrecht: Kluwer Academic.

Vellutino, F., & Scanlon, D. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly, 33,* 321–363.

Wagner, D. A. (1987). Literacy futures: Five common problems from industrializing and developing countries. In D. A. Wagner (Ed.), *The future of literacy in a changing world* (pp. 3–16). Oxford: Pergamon.

Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal, 20,* 359–373.

Walberg, H. J., & Tsai, S. (1984). Reading achievement and diminishing returns to time. *Journal of Educational Psychology, 76,* 442–451.

West, R. F., & Stanovich, K. E. (1991). The incidental acquisition of information from reading. *Psychological Science, 2,* 325–330.

West, R. F., Stanovich, K. E., & Mitchell, H. (in press). Reading in the real world and its correlates. *Reading Research Quarterly.*

White, T. G., Slater, W. H., & Graves, M. F. (1989). Yes/no method of vocabulary assessment: Valid for whom and useful for what? In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction* (38th Yearbook of the National Reading Conference, pp. 391–397). Chicago: National Reading Conference.

Wirtz, W. (1977). *On further examination.* New York: College Entrance Examination Board.

Yussen, S. R. (1990). Rethinking child development. *Contemporary Psychology, 35,* 677–678.

Zill, N., & Winglee, M. (1990). *Who reads literature?* Cabin John, MD: Seven Locks Press.

Zimmerman, J., Broder, P. K., Shaughnessy, J. J., & Underwood, B. J. (1977). A recognition test of vocabulary using signal-detection measures, and some correlates of word and nonword recognition. *Intelligence, 1,* 5–31.