9

Literacy Experiences and the Shaping of Cognition

Keith E. Stanovich, Anne E. Cunningham, and
Richard F. West

Reading is a popular topic in cognitive development and education. Within cognitive developmental psychology, for example, there is considerable literature on the individual differences in the cognitive processes that support efficient reading performance (Carr & Levy, 1990; Gough, Ehri, & Treiman, 1992; Perfetti, 1985; Share & Stanovich, 1995). A popular research strategy has been the cognitive correlates approach (see Pellegrino & Glaser, 1979; Sternberg, 1990) in which investigators attempt to determine whether individual differences in particular cognitive processes or knowledge bases can serve as predictors of reading ability (e.g., Carr & Levy, 1990; Jackson & McClelland, 1979). The causal model that is implicit in such analyses locates individual differences in the cognitive subprocesses prior to reading ability. The focus of this chapter tends to invert the causal model implied in most of this research. That is, many researchers have attempted to specify individual differences in the cognitive processes that support efficient

---

This research was supported by Grant No. 410-95-0315 from the Social Sciences and Humanities Research Council of Canada to Keith E. Stanovich; a postdoctoral fellowship from the James S. McDonnell Foundation to Anne E. Cunningham; and a James Madison University Program Faculty Assistance Grant to Richard F. West.

reading performance. In contrast, very little attention has been focused on the reciprocal possibility that exposure to print itself, *print exposure,* affects the development of cognitive processes and declarative knowledge bases.

In contrast to the relative inattention to the consequences of reading experience displayed by developmental psychologists, the literature on the cognitive consequences of literacy in the humanities and social sciences outside of psychology is large (Gee, 1988; Goody, 1977, 1987; Graff, 1986, 1987; Havelock, 1963, 1980; Kaestle, 1991; Ong, 1967, 1982; Stock, 1983). Over the past 3 decades, scholars such as Goody (1977, 1987), Olson (1977, 1994), and Ong (1982) have promulgated a view that has come to be called the Great Divide theory, which proposes that literacy fosters logical and analytic modes of thought, critical attitudes, propositional knowledge, and abstract uses of language. However, in the 1980s, the Great Divide theory received what seemed like a death blow from the much publicized study of Scribner and Cole (1981), who examined literacy effects among the Vai people in Africa. The fact that some unschooled individuals in this society were familiar with an indigenous script allowed researchers to separate schooling effects from literacy effects. Scribner and Cole (1981) found no specific effect of literacy on a number of tasks tapping general cognitive processes, including taxonomic categorization tasks, memory tasks, and syllogistic reasoning problems. The extremely innovative separation of literacy and schooling in the Scribner and Cole investigation led to an almost instant acceptance in the literature of their main conclusions on the consequences of literacy.

The seeming conclusiveness of the Scribner and Cole (1981) investigation dampened enthusiasm for new empirical studies of the effects of literacy. Unfortunately, Scribner and Cole's (1981) innovative and costly project is unlikely to be replicated, so resolving the issues using a variant of their methodology will not be possible. However, the cognitive consequences of literacy can be studied without necessarily using a cross-cultural comparison. Our methodology exploits the fact that even within a generally literate culture, individuals vary tremendously in their degree of exposure to print. Furthermore, even among a group of individuals who have the same level of assessed reading comprehension ability, remarkably large differences are found in their degrees of engagement in print-related activities (Stanovich & West, 1989), and, most important, the correlates of this natural variation can be studied. Comparing literate and illiterate people 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 the amount of exposure to print found among literate individuals. Our research program is predicated on the view that this discontinuity assumption is false, and that there is important cognitive variation among people who differ in only the *amount* of reading that they do. We do not dispute the fact that there may be important cognitive implications of the literacy–illiteracy divide, but point out that other, more continuous variability in literacy practices deserves exploration. Our research conclusions are thus restricted to the more continuous variation in reading experience.

INDIVIDUAL DIFFERENCES IN DECLARATIVE KNOWLEDGE: ALTERNATIVE VIEWS

Theories of cognitive development that have strongly emphasized the importance of declarative knowledge provide an important theoretical motivation for this research program (Alexander, 1992; Bjorklund, 1987; Ceci, 1990, 1993; Chi, 1985; Chi, Hutchinson, & Robin, 1989; Hoyer, 1987; Keil, 1984; Scribner, 1986). Given that the knowledge-dependency of cognitive functioning is a central tenet of many contemporary developmental theories, it is surprising that more attention is not directed to a question that such theories seem naturally to prompt: Where does knowledge come from? This question seems to be addressed only implicitly by theories emphasizing knowledge-dependency, the most common implication being that individuals' differences in domain knowledge are, for the most part, a product of experiential differences. In contrast, some investigators have explicitly argued against the experiential assumption implicit in the declarative knowledge literature. These alternative hypotheses can be illustrated by using vocabulary knowledge as an example.

Vocabulary is an important knowledge base for many aspects of psycholinguistic processing, and it is certainly tempting to attribute readers' variability in knowledge of vocabulary to experiential differences. For example, there is considerable evidence indicating that the size of children's vocabularies is correlated with parental education and indicators of environmental quality (Hall, Nagy, & Linn, 1984; Mercy & Steelman, 1982; Wells, 1986). Thus, it has been argued that vocabulary differences are primarily the result of differential opportunities for learning words. This conjecture might be termed the *environmental opportunity hypothesis.*

The environmental opportunity hypothesis is countered by theorists who emphasize that differences in vocabulary are caused by variation in the efficiency of the cognitive mechanisms responsible for inducing meaning from context. Proponents of what we might call the *cognitive efficiency hypothesis* argue that experiential factors are not implicated—or at least are of secondary importance—in explaining differences in size of vocabulary. 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 stated the cognitive efficiency hypothesis in even stronger form:

> 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)

It is important to realize that cognitive efficiency explanations of this type are generic and are not necessarily restricted to the domain of vocabulary acquisition. They could, in theory, apply to knowledge acquisition in virtually any domain. Ceci (1990) has discussed how in an attempt to undermine developmental theories that emphasize the importance of knowledge structures in determining intelligent performance, advocates of the cognitive efficiency hypothesis argue that "intelligent individuals do better on IQ tests because their superior central-processing mechanisms make it easier for them to glean important information and relationships from their environment" (p. 72). The cognitive efficiency hypothesis thus undercuts all developmental theories that emphasize the importance of knowledge structures in determining intelligent performance by potentially trivializing them. According to the cognitive efficiency view, these differences in individuals' knowledge bases may affect certain cognitive operations, but the knowledge differences themselves arise merely as epiphenomena of differences in the efficiency of more basic psychological processes. Differences in acquired knowledge thus become much less interesting as explanatory mechanisms of developmental differences, because they are too proximal a cause.

MEASURING THE SPECIFIC EFFECTS OF PRINT EXPOSURE

As part of a broad-based research program examining the impact of reading experience on cognitive development (Echols, West, Stanovich, & Zehr, 1996; Stanovich, 1993; Stanovich & Cunningham, 1992, 1993), we have put to the test the cognitive efficiency hypothesis by examining the experiential variable that presents perhaps the most serious challenge to it: exposure to print. Before embarking on these investigations, we were faced with two fundamental problems: (a) How do you measure individual differences in exposure to print? and (b) How should you interpret any associations between cognitive outcomes and print exposure that are observed? We turn first to the former question.

A variety of methods have been used to assess individual differences in exposure to print (Guthrie & Greaney, 1991; Smith, 1996). For example, many different questionnaire and interview techniques have been used, 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. Activity diaries yield estimates of the actual amount of time spent each day on literacy activities and are generally more valid than interview or questionnaire instruments.

Anderson and colleagues (Anderson, Wilson, & Fielding, 1988) pioneered the use of the activity diary method to estimate the amount of time that fifth graders (10–11-year-olds) spent reading in their nonschool hours, and we have used the activity diary method in some of our own studies. Our method of collecting daily activity records was adapted from that used in the Anderson et al. (1988) investigation, 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 time 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 (10–11-year-olds) sample spent in various activities when they were outside of school.

Although some of our categories were different from those of the Anderson et al. (1988) study, those that were common 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-grade students watched less television (83.2 min vs. 131.1 min) and did more homework (49.0 min vs. 18.9 min) than the Anderson et al. fifth-grade students. 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 versus 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 min, respectively) were very close to those obtained in the Anderson et al. study (10.1 and 4.6 min). Certain rough generalizations thus hold across the two studies: Fifth-grade students (10–11-year-olds) spend around 5 min per night reading books for pleasure outside of school, less than 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 have scored poorly and in which their poor performance has been linked to excessive television watching, low levels of homework, and little reading (Applebee, Langer, & Mullis, 1988; Chen & Stevenson, 1989; Stevenson & Lee, 1990; Stevenson, Stigler, Lee, Lucker, Kitamura, & Hsu, 1985).

Our specific concern, however, was to find whether the amount children read related to their achievement and whether such a linkage could be shown to have any specificity. In our study, time spent reading books (logarithmically transformed, see Allen et al., 1992, and Anderson et al., 1988) displayed a significant correlation of .39 with a standardized test of vocabulary knowledge. However, the significant zero-order correlation is not, by itself, enough to establish that vocabulary size is specifically linked to reading experience. The cognitive efficiency hypothesis is simply one way of framing the basic problem, which is that levels of print exposure are correlated with 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. Attributing any particular outcome only to print exposure is extremely difficult.

We have used a regression logic to deal with this problem. In our analyses, we first regress out general measures of cognitive ability before examining the relationship between print exposure and criterion variables. The logic of our analytic strategy is conservative because, in certain analyses we have actually partialled out variance in abilities that are likely to be developed by reading itself. However, the explanatory ambiguities surrounding a variable such as print exposure have led us to continue to structure the analyses in a so-called worst case manner, as far as print is concerned.

In this study, we assessed the specificity of the relation between reading books and development of vocabulary by conducting a hierarchical regression analysis in which a 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 scholastic learning ability. When entered second, time spent reading books explained an additional 9.7% of the variance, and this unique variance was statistically significant ($p < .01$). Thus, the linkage between vocabulary and book reading time remains even when variability in general academic performance is partialled out.

Alternative Methods for Assessing Exposure to Print

Before embarking on further tests using this logic, we needed to develop an alternative methodology for measuring print exposure that was less logistically taxing than the activity diary technique. The latter requires extensive participant cooperation over a number of weeks. Children must record their activities from the day, either at the end of the day or on the following morning, and these recordings must be checked by a teacher or another adult to assure that the scale is being used properly. Such a level of participant involvement may discourage many investigators from using the technique.

A further 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 over-report socially desirable behaviors (Furnham, 1986; Paulhus, 1984). In this case, the effect would be to report more reading than actually takes place. Independent evidence indicates that social desirability does distort self-reports of time spent reading books 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.

However, we were not constrained to use the diary method because the correlates of differential exposure to print can be studied without estimating absolute amounts of reading in terms of minutes per day. Only an index of *relative* differences is required for the regression logic to be employed. Thus, one can use measures of print exposure that do not have some of the drawbacks of the activity diary method. Our research group has attempted to develop and validate measures of individual differences in print exposure that were designed: (a) to yield estimates of relative differences in exposure to print in a single 5–10 minute session; (b) to have very simple cognitive requirements (i.e., not require retrospective time estimates); and (c) to be immune from contamination from the tendency to give socially desirable responses.

The first measures we developed (Stanovich & West, 1989) were designed for use with adult participants. 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). Participants 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, 1993; Stanovich, West, & Harrison, 1995 for examples of the stimuli).

In constructing the list of ART authors, we selected items that 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 content. Thus, an attempt was made to avoid authors who are regularly studied in the school curriculum. In short, the ART was intentionally biased toward out-of-school reading, because it was intended as an indirect measure of the amount of free reading participants engaged in.

The 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. Furthermore, 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 recalling time spent, as does the use of daily activity diaries. Finally, the measures can be administered in a matter of a few minutes.

The 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 indicators is clearly problematic in some contexts. For example, a participant's 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 participants in our studies is revealing: Some participants said they knew that a certain name was that of an author but, nevertheless, had never read anything that the author had written. When questioned about how they knew that the name was a writer, the participants 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 forth. In short, individuals' 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 (e.g., 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. 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 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.

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. Children do not need to have read the entire book or to remember any of the contents. However, 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 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.

Validation of Checklist Measures

We have validated all the recognition checklist measures in a variety of ways. First, we have shown that they are convergent with diary estimates of absolute reading time (Allen et al., 1992). In another study (West, Stanovich, & Mitchell, 1993), we attempted to validate the checklist print-exposure measures by seeing if 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 at National Airport in Washington, DC. Reading occurs in this setting by way of the free choice of the participant. If individual differences in free reading in a setting such as this could be related to performance on the recognition checklist tasks, it would bolster the construct validity of the checklist measures as indicators of individual differences in print exposure.

Individuals sitting by themselves were the potential participants and were monitored unobtrusively by the experimenter for 10 min consecutively. If participants 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. Over 90% of the potential subjects agreed to participate.

Table 1 displays the results of a comparison of the 111 readers and 106 nonreaders on a few of the checklist 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 print-exposure measures. They were reliably linked to direct observations of free reading in a situation where investigators do not intrude upon the process.

Table 1

Differences Between Readers and Nonreaders

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*

NOTE: $df = 211$ for the vocabulary checklist; 213 for the Magazine Recognition Test (MRT); 214 for film recognition; and 215 for all other variables.
$^{*}p < .01$.

The data presented in Table 1 illustrate additionally that the readers were also superior on measures of vocabulary and general knowledge (a cultural literacy test). However, as the last two rows of Table 1 show, the readers were also older and had more education. It is thus possible that age or education might have resulted in a spurious link between airport reading and performance on the vocabulary and cultural literacy measures.

The results of the two hierarchical regressions presented in Table 2 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 even after age and education had been partialled out. These regressions demonstrate that we have discovered, in essence, a "10-minute airport test" that predicts vocabulary, independent of educational level. Studies such as this and indications that the checklist measures converge with diary estimates of reading activity (Allen et al., 1992) gave us confidence in employing the

Table 2

Airport Reading as a Predictor of Vocabulary and Cultural Literacy

Step/Variable	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*

*$p < .01$.

former as measures of individual differences in print exposure in some of our other studies.

PRINT EXPOSURE AS A CONTRIBUTOR TO GROWTH IN VERBAL SKILLS

In several studies, we have attempted to link print exposure to specific cognitive outcomes after controlling for relevant general abilities—in short, to test the cognitive efficiency hypothesis. In a study of fourth-, fifth-, and sixth-grade children (Cunningham & Stanovich, 1991), we examined whether print exposure accounts for differences in vocabulary development when controls for both general and specific (i.e., vocabulary relevant) abilities were invoked. The analyses displayed in Table 3 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, Slater, & Graves, 1989); a verbal fluency measure where the children had to say as many words as they

Table 3

Unique Print Exposure Variance After Age, Performance on Raven Progressive Matrices, and Phonological Coding Were Partialled Out

Step/Variable	*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**

NOTE: The spanner headings identify the dependent variables in the regression analyses. TRT = Title Recognition Test; PPVT = Peabody Picture Vocabulary Test.
$^{*}p < .05$. $^{**}p < .01$.

could that fit into a particular category (e.g., things that are red; see Sincoff & Sternberg, 1987); and 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 control for ability that would be 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 such as 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, 1993). Thus, print exposure and vocabulary might be spuriously linked by way of their connection with decoding ability: Good decoders read a lot and have the best context available for inferring new words. 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 partialled out.

The results of the first three analyses displayed in Table 3 indicate that for each of the vocabulary measures, the TRT accounted for significant variance after the variance attributable to performance on the Raven Matrices 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 Wechsler Intelligence Scale for Children (WISC).

We have conducted an even more stringent test of whether exposure to print is a unique predictor of verbal skill in a study of college subjects (Stanovich & Cunningham, 1992). Table 4 summarizes the re-

Table 4

Unique Print Exposure Variance After Nonverbal Abilities and Reading Comprehension Ability Are Partialled Out

Step/Variable	N–D Voc.	PPVT	NAEP	CLR	Sp Com	Vb Fl
			Multiple *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*

NOTE: Dependent variables: N–D Voc. = Nelson–Denny Vocabulary; PPVT = Peabody Picture Vocabulary Test; NAEP = history and literature (from National Assessment of Education Progress); CLR = cultural literacy recognition; Sp Com = spelling composite; Vb Fl = verbal fluency; ART = Author Recognition Test; MRT = Magazine Recognition Test.

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

sults of this study. In this hierarchical forced-entry regression analysis two nonverbal measures of general ability were entered first into the equation: performance on a figural analogies test and on the Raven Matrices. Next, performance on the Nelson–Denny Reading Comprehension Test is entered subsequent to the two nonverbal ability tasks but prior to the measure of print exposure. By structuring the analyses in this way, we do not mean to imply that print exposure is not a determinant of reading comprehension ability. Indeed, we would argue that there are grounds for believing that exposure to print does facilitate growth in comprehension ability. However, in recognition of the correlational nature of our data, we have attempted to construct the most conservative analysis possible by deliberately allowing the Nelson–Denny comprehension measure to steal some variance that is rightfully attributed to the print-exposure measure. The results illustrated in Table 4 indicate that print exposure was able to account for additional variance in two measures of vocabulary, and two measures of general knowledge, spelling, and verbal fluency even after reading comprehension ability had been partialled along with nonverbal ability.

That these analyses are conservative in entering reading comprehension before the print-exposure measure is illustrated in a longitudinal study that we have conducted (Cipielewski & Stanovich, 1992), which indicates that exposure to print is related to growth in reading comprehension ability. The regression analyses presented in Table 5 display the results of this study in which growth in reading comprehension ability was tracked by administering the comprehension tests from the Stanford Diagnostic Reading Test and the Iowa Test of Basic Skills (ITBS) to 82 fifth graders who had been administered the comprehension subtest from the ITBS in the third grade (as 8–9-year-olds). The regressions are hierarchical forced-entry analyses for prediction of fifth-grade reading comprehension ability. Third-grade reading comprehension was entered first, followed by the recognition checklist measure of print exposure. Thus, the analyses are essentially addressed to the question of whether exposure to print can predict individual differences in growth in reading comprehension from third to fifth grade. In both cases, print exposure predicted variance in fifth-grade reading

Table 5

Hierarchical Regressions Predicting Fifth-Grade Reading Ability

Step/Variable	R	R^2	R^2 change	F to enter
	Fifth-grade Stanford Reading Comprehension			
1. Iowa Comprehension (third grade)	.645	.416	.416	54.06**
2. Title Recognition Test	.725	.526	.110	17.38**
	Fifth-grade Iowa Reading Comprehension			
1. Iowa Comprehension (third grade)	.545	.297	.297	33.78**
2. Title Recognition Test	.609	.371	.074	9.25**

NOTE: The spanner headings identify the dependent variables in the regression analyses.
**$p < .01$.

comprehension ability after third-grade reading comprehension scores had been partialled out. Thus, in partialling reading comprehension ability in our adult studies, we are undoubtedly removing some of the variance in the criterion variable that is rightfully attributed to print exposure.

EXPOSURE TO PRINT AND DECLARATIVE KNOWLEDGE

In other studies, we have focused even more directly on content knowledge by addressing the question "Where Does Knowledge Come From?" Stanovich and Cunningham (1993) examined general ability, print exposure, and exposure to other media sources as determinants of individual differences in content knowledge. This study contained a particularly stringent test of the cognitive efficiency explanation of individual differences in the acquisition of knowledge. The subjects were 268 college students, and the test is displayed in Table 6. The criterion variable

Table 6

Hierarchical Regression Analyses Predicting General Knowledge Composite Among 268 College Students

Step/Variable	R	R^2 change	F to enter	Final β	Final F
1. HS GPA	.372	.139	42.82**	.020	0.32
2. Raven Matrices	.447	.061	20.30**	.016	0.20
3. Mathematics test	.542	.094	35.07**	.165	18.19**
4. N–D Comp.	.630	.103	45.11**	.112	9.87**
5. Television composite	.630	.000	0.06	−.039	1.68
6. Print composite	.876	.371	417.63**	.720	417.63**

NOTE: HS GPA = high school grade-point average. N–D Comp. = Nelson–Denny Comprehension Test.
$^{**}p < .01$.

is a composite index of performance on five general knowledge measures. Four measures of general ability were entered prior to print exposure: high school grade-point average, performance on the Raven Matrices, performance on an SAT-type mathematics test, and the score on the Nelson–Denny Reading Comprehension Test. This set of tasks surely exhausts the variance attributable to any general ability construct, and general ability does account for a substantial proportion of variance in the general knowledge composite (multiple *R* of .63). When entered as the fifth step, a composite measure of exposure to television accounted for no additional variance. However, a composite index of exposure to print accounted for a substantial 37.1% of the variance when entered after the four ability measures and television exposure.

This pattern replicated in each of the five measures of general knowledge we employed, including a homemade instrument we called the Practical Knowledge Test. This task was designed to address the criticism that our other measures of general knowledge were too academic, that they tapped knowledge that was too esoteric or elitist, information that was not useful in daily life. We thought this a dubious criticism—many items on these measures were mundane and concrete questions such as "In what part of the body does the infection called pneumonia occur?", hardly esoteric. Nevertheless, in the Practical Knowledge Test, we addressed these criticisms by devising questions that were directly relevant to daily living in a technological society in the late 20th century. For example, "What does the carburetor in an automobile do?"; "If a substance is carcinogenic it means that it is _______?"; "After the Federal Reserve Board raises the prime lending rate, the interest that you will be asked to pay on a car load will generally increase/decrease/stay the same?"; "What vitamin is highly concentrated in citrus fruits?"; "When a stock exchange is in a "bear market," what is happening?"; and so forth.

The results indicate that more avid readers in our study—independent of their general abilities—knew more about how a carburetor worked, were more likely to know their U.S. senators, were more likely to know how many teaspoons are equivalent to one tablespoon, were more likely to know what a stroke was, what a closed-shop was in a factory, and so forth. One would be hard pressed to deny that at least some of this knowledge is relevant to living in the United States in the late 20th century.

In other questions asked of these same subjects, we attempted to probe areas that we thought might be characterized by *mis*information. We then attempted to trace, in our individual difference analyses, the "cognitive anatomy" of this misinformation. One such question concerned the sizes of the world's major religions and was designed to assess awareness of the multicultural nature of the modern world. The question was phrased as follows: "The 1986 *Encyclopedia Britannica* estimates that there are approximately nine hundred million people in the *world* (not just the United States) who identify themselves as Christians. How many people in the world (not just the United States) do you think identify themselves as _______?" Space was then provided on the form for the subjects to make estimates of the number of Moslem, Jewish, Buddhist, and Hindu people.

We will focus here on the estimates of Moslem and Jewish people because of our a priori hypothesis that availability effects caused by U.S.

television coverage of Israel has skewed the perception of this ratio. Although the median estimate in our sample of the number of Jewish people (20 million) was quite close to the actual figure of 18 million according to the 1990 *Universal Almanac*, the number of estimated Moslem people—a mean of 10 million—was startlingly low (817 million is the estimate in the *Universal Almanac*). For each subject, we calculated the ratio of the estimates of Moslem to Jewish people to see how many subjects were aware of the fact that the number of Moslem people is an order of magnitude larger (the actual estimated ratio is approximately 33:1 according to the *World Almanac*; 45:1 according to the *Universal Almanac*). The median ratio in our sample was 0.5. That is, 69.3% of our sample thought that there were more Jewish people in the world than Moslem people.

This level of inaccuracy is startling given that approximately 40% of our sample of 268 students were attending one of the most selective public institutions of higher education in the United States (the University of California, Berkeley). We have explored the correlates of this particular misconception in a variety of ways. In the following analysis, we scored a subject's ratio "correct" if it was 1.0 or greater—admittedly a ridiculously liberal scoring criterion, but one necessitated by the fact that only 8.2% of the sample produced a ratio of 20:1 or greater. Table 7 presents a breakdown of the scores on this question based on a median split of the print composite and television composite variables. There is a clear effect of print exposure on the scores on the question, and a significant effect of television viewing, but the effects were in

Table 7

Proportion of "Correct" Answers on the Estimates of Moslem and Jewish People as a Function of Print and Television Exposure

High print exposure		Low print exposure	
Low TV	High TV	Low TV	High TV
.614	.419	.298	.236