The idea of an attentional deficit as the central syndrome affecting learning-disabled (LD) children was popularized by Dykman, Ackerman, Clements, and Peters (1971). Since then the idea has become widely accepted by researchers in the field, both for its simple, perhaps simplistic intuitive appeal, and because of the large number of studies that claim to support it. Unfortunately, evidence for this disorder is quite difficult to substantiate, because a close look at the attention literature with LD children reveals as much disarray as does the rest of the literature with LD subjects. The problems one encounters in this literature are myriad, but they can be sorted into three general categories. There are problems with the sampling and labeling of LD subjects, problems with the definitions and measurement of attention employed, and there are a variety of methodological problems. These are discussed in turn. This chapter demonstrates that there is no support for the notion of an attentional deficit as a central syndrome affecting LD children; it also outlines new research that will substantiate this claim.

The first problem one encounters in the literature is the sampling/labeling problem. In many studies purporting to have LD children it is impossible to know who the subjects really were. Only recently have many researchers begun to separate “hyperactive” from “LD” populations, although the children described by these terms are often very separate and distinct groups. In principle, a hyperactive diagnosis implies nothing whatsoever about a child’s learning achievement or potential, whereas an LD label implies nothing about a child’s activity. And although the DSM III describes an “Attentional Deficit Disorder with Hyperactivity” (ADDH) syndrome for some hyperactive children, there is no evidence for global attentional deficits in LD populations, as this chapter demonstrates.
Another major sampling error involves the misuse or nonuse of intelligence test scores. Although children labeled LD should have normal range IQs of 85 and above, Kirk and Elkins (1975) found in a survey of over 3,000 LD children that the median IQ was 93, and that 35% had IQs below 90. Simple calculations reveal that only 9% of children in the general population would be expected to have IQs in the 85 to 90 range. Even more troubling was the finding of Keogh, Major-Kingsley, Omori-Gordon, and Reid (1982), who did an exhaustive computer search of experimental studies with LD populations. They located more than 4,000 such studies and found that more than 30% of the studies they sampled did not report IQ scores at all. The possibility is distinct, therefore, that LD/non-LD performance differences in many studies may be caused by sampling errors involving the inclusion of individuals with low IQs, as well as hyperactive children.

The second problem with the LD attention literature has to do with the notion of attention itself, and with what constitutes evidence for a deficit in attention. Faulty attention, in the most global sense of the term, can explain poor performance on any task. Such a notion is therefore useless as an experimental hypothesis. Some researchers prefer to explore one or more of the “components” of attention in order to provide support for the attentional deficit notion. The components studied vary somewhat, but the trend is to include “arousal,” “vigilance,” and “selective attention” as three of the most common components (e.g., Samuels & Edwall, 1981; Tarver & Hallahan, 1974). Typically, each group of researchers devises one experiment with one task on which LD/non-LD differences are demonstrated. It can certainly be argued, as does Koppel (1979), that poor performance on one task designed to measure one component of attention does not substantiate the claim of across-the-board attentional deficits. In fact, such a claim could only be supported with evidence, using one carefully selected LD population, that showed both poor LD performance on a number of attention tasks and significant correlations among performance scores on the various tasks. Without such correlations, one cannot conclude that the various experimental tasks, given to separate LD groups, are tapping the same “thing” called attention at all, regardless of what the various researchers may claim to be measuring. Nevertheless, disparate studies that show LD/non-LD differences on diverse tasks, measuring different components of attention, with different research populations, are taken collectively to support the claim of global attentional deficits in LD children (e.g., Bryan & Bryan, 1980; Hallahan, 1975; Keogh & Margolis, 1976). Such claims are simply not warranted. No researcher has demonstrated correlations among performance scores on various attention tasks with a single, nonhyperactive LD population. Pelham (1979) did administer multiple attention tasks to a single group of subjects. Notably, he did not find clear-cut attentional differences between LD and non-LD children, nor did he find correlations among performance scores on his tasks. This study is discussed in more detail in a later section.
The final sort of problems one encounters in this literature are methodological. Although individual studies exhibit a variety of problems, there are some problems of a more general nature that seem to be epidemic in the LD attention literature. One quite serious problem is that rarely, if ever, are baseline differences in performance between the LD/non-LD groups equated; another is that group IQ differences are almost never eliminated or controlled (see discussion in Douglas & Peters, 1979). Moreover, the attention tasks used in many experiments are frequently heavily memory-loaded or otherwise strategy-dependent, such that any poor student would be expected to perform poorly. Finally, the dependent measure of attention is often reaction time (RT). Whereas RT may be a fair approximation of the attending behavior of a college sophomore, it may not be a realistic measure of the capabilities of an unmotivated youngster. These problems are elaborated in the sections that follow.

THE COMPONENTS OF ATTENTION

Arousal

It has been suggested by many psychologists (e.g., Kahneman, 1973) that there are optimal levels of physiological arousal that correspond with the difficulty levels of various tasks. A subject’s attention is inferred from his physiological state, which is measured by reaction time, heart rate, EEG, pupil size, and so forth. It is reasoned by some that if LD children are less reactive than non-LD children in performing certain tasks, then they are exhibiting attentional disorders. (See Rosenthal & Allen, 1978, for a review.) LD researchers who study the arousal component of attention often work in the psychiatric tradition. A major problem with much of this research is that the LD subjects who participate in these studies are usually clinical referrals to medical school research laboratories. The children often have neurological soft signs and very unacceptable classroom behavior. Hyperactive children are routinely included in the research samples. The researchers typically record physiologic measures and reaction times while the children perform complicated tasks. Not surprisingly, LD children are often found to perform poorly in such circumstances (e.g., Dykman, Ackerman, & Oglesby, 1979).

Among the researchers who work in the psychiatric tradition, reaction time is a very common dependent measure. Rosenthal and Allen (1978) reported that RTs of LD children are usually found to be slower than those of non-LD children, that greater individual variability is often found, and that there is a greater likelihood that the RTs of LD youngsters will increase over time in continuous performance tasks. This latter finding suggests that some factor other than a unified attentional dysfunction may be affecting response time. One cannot help but wonder whether a very interesting reward (say, brief glimpses at pictures of
nudes) would eliminate possible group differences in reaction time, in these failure-oriented LD youngsters.

The EEG has been another measure employed by those who work with clinical populations. However, in a major review of the neurologic tests given to these children, Coles (1978) concluded that neither the EEG nor any of the tests in the standard neurologic battery could conclusively discriminate LD from non-LD children. In practice though, very few LD children are ever given these tests because very few children are referred to medical schools. Moreover, even if LD/non-LD differences on these tests are sometimes found in such populations, it would not be safe to generalize from such findings to the average LD child. Yet, this is often done.

Vigilance

Vigilance, or sustained attention, is another component common to many theories of attention. Learning-disabled children generally perform poorly on tasks designed to measure this. A typical task consists of isolating the child for 20 to 30 minutes with a panel of flashing lights or a rotating drum displaying letters. The child must make a rapid response to a particular pattern of lights or sequence of letters. The poor performance of LD children on these tasks is commonly offered as evidence of an attentional deficit (e.g., Anderson, Halcomb, & Doyle, 1973). One could certainly make the case, however, that many factors—interest, motivation, desire to conform, and so forth—can contribute substantially to a child's score on a vigilance task. It is reasonable then to suggest that a significant portion of what is measured on such tasks may be things other than attention. Such tasks do measure a behavior pattern that is related to academic success, however, because obedience and willingness to persevere in tedium will certainly help a child succeed in school. But one must certainly question whether failure-oriented LD children would be likely to demonstrate their optimal performance capabilities in such dull, repetitive tasks.

Thus, although sustained attention tasks may frequently yield LD/non-LD differences, it may be that the tasks are not appropriate for the measurement of attention with these subjects. Such tasks may merely highlight obvious motivational differences between groups, rather than illustrating any differences in actual attentional capabilities. Such studies tell us very little about the optimal performance of LD children, because they do not provide optimal circumstances in which the children might demonstrate their abilities. It seems reasonable to suggest, therefore, that the results of both arousal studies with clinical LD populations and tedious vigilance tasks with unmotivated, atypical youngsters cannot be taken as very strong support for the notion of attentional deficits in LD children.
Selective Attention

Tasks designed to measure the selective component of attention are the most common in developmental psychology. LD/non-LD performance differences on these tasks are offered as the best evidence of the attentional deficit syndrome. An unquestioned assumption among those who use these tasks is the belief that these tasks are measuring the same unitary component, factor, or ability. This ability is variously called selective attention, distractibility, focusing, discriminability, or stimulus selection, with selective attention the most common term.

There are two general types of experimental paradigms that are used to measure the selective attention deficit. LD children appear to have trouble in tasks with proximal or embedded kinds of distractors in which the extraneous or incidental information appears to affect their performance on a central task. And they also appear to perform more poorly than non-LD children on tasks which measure the central and incidental information directly; that is, LD children are said to pay more attention to the incidental information and less attention to the central information than non-LD children. They are thought to be deficient at selective attention because they inefficiently trade-off some central information for the incidental (e.g., Ross, 1976).

The Proximal/Embedded Tasks

Stroop Tasks. It is frequently claimed that LD children have trouble with stroop tasks. The classic color Stroop (Stroop, 1935) requires the subject to name the ink color of color words that are printed in discrepant ink hues. A closer look at the research that is cited, however, reveals that the results are somewhat ambiguous. Silverman, Davids, and Andrews (1963) is frequently cited, although an examination of their data shows that when they corrected for base rate differences in naming speed, a significant difference in distractibility did not occur for their "underachievers" until the fifth trial. A trial consisted of reading a color interference card with 100 words on it. It is reasonable to suppose that the underachievers may have lost interest in the task by the 400th word or so. Eakins and Douglas (1971) claimed to show differences between poor and good readers, but they failed to control for group differences in baseline naming speed. If they had, it appears that their effect would have disappeared.

A nonreading version of the Stroop task, "The Fruit Distraction Task," or Fruit Stroop, has also been used (e.g., Santostephano, Rutledge, & Randall, 1965). The distraction condition in this task involves naming the real-world color of inappropriately colored fruit pictures, such as a purple banana. Studies with this paradigm have generally used "poor readers" (not necessarily with normal IQs), and results have been mixed. A variety of conflicting results are reported by Alwitt (1966), Campbell (1969), Denney (1974), and Santostephano et al. (1965). It is still unsettled how a well-selected sample of LD children would
perform on this task, although one would expect them to perform more poorly than controls if they evidenced a deficit in selective attention.

*The Picture–Word Stroop.* Another type of Stroop-like task involves words embedded in pictures. The child's task is to ignore the embedded words and to name the pictures as fast as possible. It has been shown that children as young as late first grade show Stroop-like interference on this task, and that naming latencies increase, in order, as nonsense words, words unrelated to the pictures, and words from categories related to the pictures are embedded; that is, interference increases as the distractor words become more semantically related to the pictures. The same pattern of interference is found for adults, although overall latencies tend to decrease with age (Ehri, 1977; Golinkoff & Rosinski, 1976; Guttentag & Haith, 1978; Rosinski, 1977; Rosinski, Golinkoff, & Kukish, 1975).

A few of these studies have used "poor readers" and have found the same semantically related interference effects, although the overall response latencies are generally longer than for the controls. However, the poor readers in these studies seem to be never more than 1½ years below grade level in reading (Golinkoff & Rosinski, 1976; Pace & Golinkoff, 1976). Frequently, no reading scores at all are reported for the poor readers (Ehri, 1976; Ehri & Wilce, 1979; Guttentag & Haith, 1978). Because we cannot safely assume that all these poor readers had normal range IQs, it would be interesting to administer this task to a carefully selected sample of learning-disabled children. Given that LD children might have slower baseline picture-naming latencies (RTs), if they were to exhibit a selective attention deficit in performing this task, then the amount of interference they experience would be greater than overall slower baseline picture-naming speed would predict.

*The Bender Test.* Some versions of the Bender Visual Motor Gestalt Test (Bender, 1946) have also been used to support the notion of the selective attention deficit. The test requires children to replicate nine simple, geometric designs. The Bender has been popular with reading specialists for years, although many now believe that it is not particularly good at discriminating poor readers from good ones (Coles, 1978). There is some evidence though that when the Bender test is made more distracting, LD children have trouble with it. Sabatino and Ysseldyke (1972) drew extraneous backgrounds around the simple Bender figures and found that a sample of 143 nonreading third graders with normal range IQs performed significantly worse than a group of matched controls. Their performance with the standard Bender figures was equivalent to that of the controls, however.

Other researchers have worked with the standardized *Background Interference Procedure* (BIP) for the Bender test (Canter, 1966). In this test the extraneous background stimuli are on the sheet of paper onto which the subject copies the
3. IN SEARCH OF ATTENTIONAL DEFICIT

Bender designs. The subject copies the figures twice, once onto a white sheet of paper, once onto a sheet with wavy, intersecting BIP lines on it. The difference between the score on the standard Bender test and the score on the BIP paper is the measure of distraction, or selective attention efficiency/inefficiency.

Fabian and Jacobs (1981) found that their sample of 14- to 18-year-old LD adolescents performed significantly worse in the BIP condition than did the normal 8- to 12-year-olds in the BIP manual tables. Other studies with this test have used populations that were more clinical. There is some evidence that this test can identify “hyperactive” children (Adams, Hayden, & Canter, 1974); children with “cerebral dysfunction” (Kenny, 1971); and “minimally brain damaged children” (Hayden, Talmadge, Hall, & Schiff, 1970). Because all these labels are frequently used to describe LD children, one might presume that LD children would perform poorly on this test if they are more distractible, and suffer from a selective attention deficit.

The Embedded Figures Test. Results of studies with the Embedded Figures Test (EFT) and the Children’s Embedded Figures Test (CEFT) (Witkin, Oltman, Raskin, & Karp, 1971) have also been used as support for the notion of a selective attention deficit. These tests are said to measure a relatively stable perceptual style known as “field dependence/independence.” The task is to find and outline a simple geometric form that is obscured in an embedding background of lines and shading. LD children are often found to be field dependent on these tests, and field dependent children are thought to have trouble with stimulus selection. Many LD researchers consider this to be evidence of a deficit in selective attention.

Results of experiments with LD subjects have been mixed, however. Both Goldman and Hardin (1982) with the CEFT and Spellacy and Peter (1978) with the EFT found LD children performed significantly worse than non-LDs. Scott and Moore (1980) found no LD/non-LD group differences on the CEFT. Tarver and Maggiore (1979) found no LD/non-LD group differences either. Cowen and Harway (1976) also found no LD/non-LD group differences, but a subgroup of their LD sample with “visual/perceptual problems,” as measured on subtests of the Illinois Test of Psycholinguistic Abilities, performed significantly worse than non-LDs on the CEFT. There does appear to be a trend for LD/non-LD group differences. However, it must be noted that the EFT is known to correlate moderately with IQ (Witkin et al., 1971), so there is some possibility that IQ differences between groups may have contributed to group performance differences in some of the studies cited.

Selective Listening. LD children are also said to demonstrate an attentional deficit in selective listening tasks. The child’s task is to select and repeat one of two different auditory messages that are presented simultaneously. Mulberg and Whitman (1978) found that reading-disabled (normal IQ) children performed
poorly on this task. The children made more intrusions from the unattended ear and more errors shadowing, or repeating, the selected message. On the other hand, Pelham (1979) found that a group of reading-disabled children did not experience more overall distraction in a selective listening task, but they had problems only in the speeded conditions of the task. Lasky and Tobin (1973) found a group of 11 “pre-LD” first graders to be not distracted from following the instructions on an audio tape by a white noise distractor, but they were more distracted than controls by a tape of a competing linguistic message.

It is not unreasonable, however, to expect that poor students would perform poorly on complicated, unnatural tasks like selective listening. Bryan and Bryan (1980) argue that many extraneous factors may contribute to a child’s poor performance on auditory tasks: the child’s age, the complexity of the task, the type of task, the responses demanded, and the child’s verbal skills. All of these may interfere with the researchers’ ability to measure attention in selective-listening tasks with young, atypical subjects.

The Central/Incidental Studies

The second type of studies that are cited as evidence for the selective attention deficit involve paradigms that measure the central and incidental information directly. If one conceptualizes selective attention as a trade off between central and incidental information, as many researchers do, and if LD youngsters have a deficit in selective attention, then they should pay relatively more attention to incidental information and less attention to central information than non-LD children. Higher incidental recall by LD children has been found in two studies. Diekel and Friedman (1976) gave non-LD and LD children a difficult IBM card-sorting task. They found that the LD children made more errors in sorting the cards (the central task) and that they also remembered more extraneous markings on the cards (the incidental measure). Likewise, Mondani and Tutko (1969), in a widely cited study, found underachievers more likely to notice and recall extraneous doodles on a questionnaire. There was no measure of central recall in this study, however.

In both of these studies, although the amount of incidental information recalled by the LD groups was greater than that recalled by the controls, it was not demonstrated that the LD children were unable to select the central information. It might have been the case that the card sorting was difficult to do, and the questionnaire was difficult to read, so the LD children directed their attention to less aversive incidental stimuli. This is not a demonstration of attentional deficit, because in neither case was there a baseline measure of performance on the central task without the distractors. If LDs are trading central for incidental information, it is crucial to measure their performance on the central task without the distractors, so that one can measure whether their central performance is impaired by the addition of the incidental distractors. And even if the central performance is impaired, the burden of proof would be on the researcher to show that the child can not, as opposed to will not, attend to the central display. This
criticism can be leveled against many of the studies that purport to show selective attention deficits.

The Hagen Paradigm. The most widely cited evidence for the selective attention deficit in LD children comes from studies with Hagen’s (1967) Incidental Learning Paradigm. The task requires that a child remember the positions of animal pictures in a repeatedly presented, random, seven-card array, and that the child later recall the incidental pictures that had been paired with the animals. Critiques of this paradigm have been put forward by many researchers, notably Douglas and Peters (1979) and Lane (1980). Nevertheless, results of studies with this paradigm are still presented as evidence for the selective attention deficit. Suffice it to say that Hagen himself claims that strategies such as verbal rehearsal confound the measurement of attention with this task (Hagen & Kail, 1975). Furthermore, when LD children are taught verbal rehearsal strategies their performance approximates that of non-LDs (e.g., Tarver, Hallahan, Kauffman, & Ball, 1976).

If one were to allow that the Hagen task measured selective attention, and one wished to substantiate an LD deficiency, it would be necessary to show that LD children pay relatively more attention than do non-LDs to the incidental information on this task. A variety of researchers have repeatedly used this task and they have never been able to demonstrate this: Copeland and Wisniewski (1981); Dawson, Hallahan, Reeve, and Ball (1980); Hallahan, Gajar, Cohen, and Tarver (1978); Hallahan, Kauffman, and Ball (1973); Heins, Hallahan, Tarver, and Kauffman (1976); Mercer, Cullinan, Hallahan, and LaFleur (1975); Pelham (1979); Tarver, Hallahan, Cohen, and Kauffman (1977); Tarver et al. (1976); Tarver and Maggiore (1979); Quay and Weld (1980). Incidental recall scores are never found to be significantly different between groups; in fact, they are often lower for the LD groups. Moreover, the correlations between central and incidental recall almost never appear to differ significantly. Memory for the central stimuli is the only real difference that is consistently found. One exception was a study by Pelham and Ross (1977), which did find incidental recall to be higher for the LD group, but the researchers used a nonstandard administration of the task.

Surprisingly, all the aforementioned studies are cited as evidence for the alleged selective attention deficit of LD children. In order to provide some support for this claim, Hallahan created the “Index of Selective Attention Efficiency”: the percentage of central recall minus the percentage of incidental recall (%C-%I; Hallahan, 1975). Nevertheless, significant %C-%I group differences are reported in only two of these studies (Copeland & Wisniewski, 1981; Hallahan et al., 1978), and trends are reported for only two others (Dawson et al., 1980; Tarver et al., 1977).

Yet another problem with this literature is that there is never a baseline measure of central recall without the incidental distractor pictures. Therefore, given the typical finding of equal incidental recall scores between LD/non-LD groups, one
cannot assume that the central recall of LD children has been differentially affected by the distractor pictures at all. Rather, one can only safely assume that LD and non-LD children attend about equally to the incidental information, and that they may differ in their purposeful central recall strategies.

**SUMMARY AND NEW EVIDENCE AGAINST A SELECTIVE ATTENTION DEFICIT**

Although results of studies with the Hagen paradigm have been the most commonly presented evidence for the alleged selective attention deficit of LD children, such studies have consistently demonstrated only lower central memory scores for LD children. The other type of experiments that have been offered as support for the selective attention deficit are experiments in which recall for incidental, proximal or embedded distractors is not measured directly, although the presence of these distractors seems to disrupt the central performance of LD children more so than non-LD children. In this tradition there is some evidence that LD children are more distracted by Stroop interference, by embedded figures, by extraneous backgrounds on Bender figures, and on several experimenter-designed tasks that have these sorts of properties. Admittedly, there are problems with some of these studies, and the results have been mixed in several cases, but if there is such a thing as a selective attention deficit it should show up on several tests designed to measure this deficit convergently. Hagen and Kail (1975) claim that performance on these sorts of tasks should correlate, because these tasks are assumed to measure a common ability. And, as was discussed earlier, such correlations are necessary in order to demonstrate that the tasks actually measure the same thing, a thing that we call attention.

One researcher who examined correlations among performance scores on several selective attention tasks was Pelham (1979). As was mentioned previously, he did not find clear-cut LD/non-LD differences on four attentional tasks, and he did not find significant correlations among the tasks. His task were in different modalities, however, and this may have attenuated some of his correlations. Moreover, results of two of his four tasks might have been interpreted as supporting the notion of a selective attention deficit, although he argued against such an interpretation. Finally, an examination of his data reveals that his task reliabilities were low \(M = .32\), and that two of his six correlations were not significantly different from their theoretical ceilings. Such problems make clear interpretations of his data difficult.

One must conclude, therefore, that the evidence for a selective attention deficit in LD children is certainly less than satisfactory, although the issue is not entirely resolved. Unambiguous empirical testing of such a notion has not previously been done. Moreover, the very notion of selective attention as a unitary factor or ability on which there are stable, individual differences can also be challenged.
Researchers who study the development of selective attention find that the age at which selective attention appears to be efficient depends on a number of variables including which tasks the subjects are given, and how the tasks are explained to them (e.g., Sexton & Geffen, 1979). Eleanor Gibson has repeatedly argued that infants are selective from birth, and that selection efficiency depends upon task appropriate skills and expectancies (e.g., Gibson & Rader, 1979). It is reasonable, therefore, to suppose that diverse selective attention tasks do not uniformly measure a thing called selective attention, but that different tasks require different specific strategies, prior knowledge, coordination, and skills, all of which affect a child's performance on these tasks. Sensible research hypotheses, therefore, are: (a) that systematic LD/non-LD differences in selective attention performance cannot be demonstrated, and (b) that different selective attention tasks do not measure a unitary ability.

A test of these hypotheses (McNellis, 1984) was recently completed. Five selective attention tasks were employed, all in the visual modality, all with proximal or embedded distractors. On all these tasks previous researchers had shown some evidence for LD/non-LD performance differences. Included were a Stroop Color-Word task, a picture-word Stroop task with four types of distractors, a modified version of the Group Embedded Figures Test, and the Bender Background Interference procedure. There was also an experimenter-designed, central/incidental task similar to the Hagen task but modified to decrease the memory load and make it less strategy dependent. In this task, the "picture-picture Stroop task," the child had to name rapidly seven animal pictures, which were repeatedly presented in random order, and to ignore seven pictures of household objects, which were embedded in or proximal to the central animal pictures. As a separate measure, the children were later asked to recall the central and the incidental pictures.

Each task was designed to have a "baseline" and at least one "distractor" condition. In the case of the three Stroop-like tasks, the baseline conditions involved rapidly naming the colors or pictures with no distractors present. In the Bender and the modified Embedded Figures Test the baseline measures were the number of errors that the child made in finding or reproducing the figures without extraneous background lines or shading. In this manner, it was possible to test whether performance in the baseline condition was significantly different from performance in the distractor condition. More important, the effects of possible group differences in performance in the baseline condition could be statistically removed from the distractor condition to yield a more precise measure of "selective attention." For example, although the Stroop-like tasks had reaction time measures, the effects of possible slower baseline RTs in the LD group could be removed.

It was predicted that although LD children might perform more poorly than controls in the raw baseline and distractor conditions of the five tasks, they would not perform more poorly on the selective attention measures, when the effects
of differences in baseline performance were removed. Also, it was hypothesized that performance in the baseline conditions would correlate, and might correlate with reading ability. However, selective attention measures would be uncorrelated, and would not correlate with reading ability. This would demonstrate that the children were performing in a systematic manner on the tasks themselves, but that only the selective attention measures of the tasks were uncorrelated. Finally, it was predicted that LD children would not trade-off central information for incidental in the recall measure of the picture–picture Stroop task; that is, they would not exhibit higher incidental memory scores at the expense of lower central memory scores.

Sixty-seven children, 36 LOs and 31 non-LO controls, participated in the study. IQ scores, and scores on standardized reading tests administered within approximately 1 month of the experiment were available for all LD children. The LD children were selected to have normal-range IQs of 84 and above ($M = 98$, $SD = 10.91$). Their reading scores averaged 1.9 years below grade level. Because IQ tests are not given to normally achieving children, reading level was used as a fair approximation of normal-range IQ for the nonLD children. The 31 non-LD children had, on the average, reading scores that were within 1 month of their grade level. Parental occupation information was available for most of the children and this was used as a rough measure of socioeconomic status (Davis, Smith, & Stephenson, 1980). The children were matched as closely as possible in age, IQ or achievement, SES, and race. The groups differed significantly in sex, however, and this was taken into account in the data analysis, as was the greater variability in age of the LD group.

The results of the experiment showed that the tasks were sufficiently distracting for the children. An analysis of the raw data showed that the distracting (D) conditions were significantly more difficult than the nondistracting (ND) baseline conditions for all five tasks for the nonLO children, and for four of the five tasks for the LD children. Moreover, as was predicted, LD children performed more poorly than non-LDs in both the raw baseline and distractor measures on four of the five tasks.

As was mentioned previously, it was necessary to devise a method whereby the effects of possible group differences in performance in the baseline conditions could be statistically removed from the distractor conditions. This would provide the measure of selective attention. Simple D minus ND difference scores, or D/ND ratios could not be used because such measures are inherently unreliable, because they assume equal, linear relationships between D and ND for all tasks, and because correlations among variables transformed in this manner may be spurious (Cohen & Cohen, 1983, chaps. 2, 10).

Therefore, the data from all 67 children were combined and multiple regression equations were computed for each of the five tasks. For each task possible baseline, sex, and age differences in performance were removed from performance in the distractor condition. The residuals from these regression lines (the
Residualized Gain Scores) were the data of comparison, the measure of selective attention. Thus, the selective attention measures were that portion of the distractor score that was not predicted by baseline performance, age, or sex. (This method is discussed in Cronbach & Furby, 1970.)

The first major hypothesis of the investigation was that LD/non-LD differences would not be demonstrated on multiple selective attention tasks. To test this hypothesis, t tests were used to measure the mean difference of the residuals, the selective attention measures, between LD/non-LD groups. This procedure is directly analogous to analysis of covariance, with sex, age, and baseline performance as the covariates, and the distraction scores as the dependent variable. This method is more useful than analysis of covariance, however, because the regression equations yield residuals, which can be used for group comparisons, for power calculations, for correlations, for reliability calculations, and for computing theoretical ceilings on the observed correlations. Results of the t tests showed no significant differences between the groups on any of the five selective attention measures, although the Bender BIP approached significance, \( t(64) = 1.82, p < .08 \), two-tailed. The average p-value for the tests of group differences on the five tasks was .48.

One can safely assume that no real population differences existed between the LD and non-LD groups on the selective attention measures of the five tasks, because the power analyses demonstrated that there was sufficient statistical power to detect differences on any task, had any differences existed. The average power to detect a mean group difference as small as 15% on any task was found to be .80. All tests were thus sufficiently powerful to detect a difference, yet none was found.

The second major prediction of the study was that selective attention would not appear to be a stable, unitary factor or ability. If the five tasks are measuring the same ability, as many developmental psychologists have suggested, then performance on the selective attention measures should correlate. Correlations were computed among the five selective attention measures and only 1 of the 10 intercorrelations was significant, with an uncorrected \( \alpha \) of .05. The average correlation among the tasks was only .11. Moreover, the Omnibus Null Hypothesis, that all rs might really be zero, could not be rejected, \( \chi^2(10) = 13.91, .10 < p < .25 \) (Cohen & Cohen, 1983, p. 58).

To ensure that unreliability in the five selective attention tasks did not cause the correlations to be so low, split-half or odd–even reliabilities were computed for each of the five tasks. These reliabilities were used to compute the theoretical ceilings for each of the 10 intercorrelations. Each correlation was then compared to its theoretical ceiling, and all were found to be significantly below their ceilings, at least \( p < .01 \), in all cases. One can conclude, therefore, that the low intercorrelations among the tasks were not caused by lack of reliability in the selective attention tasks, rather, that the tasks do not appear to be measuring a unitary factor or ability. Additional support for this assertion comes from an
exploratory Principal Components Analysis, which indicated that five separate factors are required to explain the data. These findings cast serious doubt on the notion of selective attention as a unitary ability on which there are stable individual differences.

Performance in the baseline conditions was examined next. This was done in order to ensure that the subjects were not performing randomly in all task conditions, but that only the selective attention measures were uncorrelated. Five new regression equations were generated, which removed the effects of differences in age and sex from the baseline scores. This allowed for direct comparison with the selective attention measures. It was predicted that the baseline measures would intercorrelate, although selective attention measures would not, and that baseline measures might correlate with reading ability, although selective attention measures would not.

These hypotheses were confirmed. Performance in the baseline condition proved to be quite systematic. The average of the 10 intercorrelations among baseline measures was .35, and 7 out of 10 were significant. The Group Embedded Figures Test accounted for the 3 nonsignificant correlations, as it appeared only to correlate with the Bender BIP test. Recall that the average correlation among the selective attention measures was only .11, however. Correlations between baseline performance and reading level also provided proof that the data were systematic. The average correlation of reading level with the 5 baseline measures was $-0.23$, with 3 out of 5 significant. The mean correlation of selective attention and reading ability was only $-0.08$, however, and none were significant. (These correlations are negative because the dependent measures in the five tasks are errors and latencies.)

In addition to the predictions about the five selective attention measures, there was an additional prediction about the central (C) and incidental (I) recall measures on the picture–picture Stroop task. Unlike the Hagen task, the recall measures on this task could not be substantially affected by strategy or rehearsal, because the task was primarily a Stroop task. Therefore, it was predicted that although LD children might recall somewhat less central information, they would not recall more incidental information than the non-LDs; that is, they would not trade-off central information for incidental. An analysis of the raw recall scores showed that, surprisingly, LD children recalled slightly more central and incidental information than non-LDs. These differences were not significant, however, when effects of differences in age and sex were removed.

Another measure that was used to compare LD/non-LD groups on the recall measure of this task was Hallahan's "Index of Selective Attention Efficiency," $\%C-\%I$. This difference score was computed for the raw central and incidental recall data for the LDs and the non-LDs. If this index measures selective attention efficiency, and if non-LDs are more efficient than LDs, then one would expect this index to yield group differences. No significant differences between groups were found. Moreover, if $\%C-\%I$ can be taken as a measure of selective attention,
one might expect it to correlate with the five other selective attention measures, the residuals for the five tasks. This analysis was also done. The average correlation of $\%C-\%I$, after the effects of age and sex differences were removed, with the other selective attention measures was .05. None of the correlations were significant. Thus, there is convincing data to support the position that LD children do not appear to differ from non-LD children in either central or incidental recall, when strategy variables such as rehearsal are not confounding the task. The final prediction of the study was confirmed.

CONCLUSIONS AND IMPLICATIONS

This chapter has been an attempt to outline the major evidence for the central syndrome that is said to affect LD children, the attentional deficit disorder. It was argued that evidence from studies with the arousal and vigilance components of attention does not provide clear-cut support for the attentional deficit notion, because such measures are confounded with many other variables, such as motivation, that hamper the measurement of attention with young, atypical subjects. Next, the evidence from selective attention studies was reviewed in some detail. These studies appeared to indicate a tendency for poorer LD performance on a variety of tasks, although some methodological problems with these studies were discussed. Critically, none of the studies reviewed, with one exception, had even attempted to question whether these various attention tasks were really measuring the same “thing” called selective attention at all.

Finally, some recent research with LD children (McNellis, 1984) was outlined. No evidence was found for the alleged selective attention deficit in learning-disabled children in this study, although evidence from studies with the selective component of attention had previously proved to be the best evidence for the alleged attentional deficit disorder of LD children. Five selective attention tasks were used in this experiment, and on only one task, the Bender BIP, was there even a tendency for LD children to demonstrate more distraction than the non-LDs. Each test had enough statistical power to detect group differences, yet no differences emerged. A sixth measure of central and incidental recall, which has often been used as a measure of selective attention, also failed to demonstrate significant group differences. These central and incidental recall scores were examined several different ways, and no measure yielded significant LD/non-LD differences.

On all these tasks, or tasks very similar to them, other researchers have sometimes demonstrated LD/non-LD differences. One reason that group differences were not found in this study may be because the subjects were selected and matched on IQ or achievement, SES, race, sex, grade, and age as carefully as practical restraints in real-world research would allow. The greater age variability of the LD children was taken into account in the data analysis, as was
the inevitable sex difference between groups. Finally, baseline performance differences between the groups were controlled. This was particularly important because three of the tasks required a reaction time response. Other researchers who use selective attention measures with LD populations rarely match groups on all these variables, nor do they routinely use multiple regression to control for differences that may exist. Other researchers may have found group differences on similar tasks because their samples may not have been as carefully selected; because their tasks, notably the Hagen task, may have been confounded by other cognitive variables; or because they may not have controlled for baseline response differences between groups. One can, however, assert with a fair degree of confidence that the reason that no LD/non-LD differences were found on these five selective attention measures is that no population differences exist.

The second major finding of the study was that there were no significant correlations among the selective attention measures, although baseline performance scores for the tasks were correlated. The tasks had been designed to be highly similar to one another to increase the chance of intercorrelation. All tasks were visual, and all required attending to a central visual stimulus in the face of proximal or embedded distractors. Furthermore, the low observed correlations were not caused by lack of reliability in the data. All observed correlations were significantly and substantially below their theoretical ceilings. It appeared that the five tasks were not measuring a unitary selective attention ability, because performance on the selective attention measures was uncorrelated.

The major implications of the study concern the nature of selective attention itself and its relationship with developmental psychology. The results of the study indicate that developmental psychologists have been willy-nilly researching an “ability” that may not exist. Neither these data, nor those of Pelham (1979), who used a different battery of tests and a different subject pool, were able to demonstrate correlations in performance among several selective attention measures. Previous developmental researchers have each used a single type of selective attention measure, and their findings are as diverse as their tasks and instructions. Psychologists and educators are going to have to look somewhere other than at “attention” to find a single syndrome that may typify learning-disabled children.

ACKNOWLEDGMENT

This study was supported in part by the Cornell Comprehensive Cognitive Psychology Training Program, under NIMH Grant 5T32 MH15777.

REFERENCES

3. IN SEARCH OF ATTENTIONAL DEFICIT


