Effectiveness of a Sensory Integrative Therapy Program for Children with Perceptual-Motor Deficits

Julie F. Densem, Graham A. Nuthall, John Bushnell, and Jacqueline Horn

This study was an evaluation of the sensory integrative therapy (SIT) program (Ayres, 1972a) for children at the Christchurch Hospital. Fifty-five children were randomly assigned to the SIT program, a parallel physical education program, or to a no-treatment condition. The children were assessed before and after treatment on measures of perceptual-motor development, language and reading development, self-concept, and handwriting skills. Covariance analysis, with age and pretest scores as covariates, found no significant differences between groups on any of the measures except reading progress among those children who could already read at the beginning of the program. Children who made the least progress during therapy were those who (a) had epilepsy, (b) were from a low-income, single-parent family, or (c) had behavioral problems.

Sensory integrative therapy (SIT) (Ayres, 1972a) is a controversial perceptual-motor training program designed for children with learning disabilities (LD) who exhibit poor sensory-motor development (Bochner, 1980; Goodman & Hammill, 1973; Jenkins & Sells, 1984; Sieben, 1977). It is normally implemented by occupational therapists.

Sensory integrative therapy follows the neuropsychological process approach to remediating learning disabilities. Unlike programs that work directly on the skills that define the learning problems, SIT involves activities designed to remediate perceptual and motor deficits believed to underlie and cause the learning problems. This neuropsychological process approach can be divided into several different perceptual-motor theories (Ayres, 1972a; Barsch, 1967; Cratty & Marlin, 1969; Frostig, 1967; Getman, Kane, Halgren, & McKeen, 1968; Kephart, 1960) and the neurological organization theory (Doman, Spitz, Zucman, Delacato, & Doman, 1960). Sensory integrative therapy is based on a perceptual-motor theory, but differs from other therapies based on perceptual-motor theory in its greater emphasis on central nervous system etiology and less emphasis on visual skills training.

Reviewers of the research on SIT have come to very different conclusions about its efficacy. Bochner (1980) and Green, Hope, Oates, Parry, and Procopis (1982) decided that the evidence was inconclusive, while Ottenbacher (1982) concluded that SIT does have empirical support. There is reason to believe, however, that Ottenbacher's conclusions were based on a number of studies that had serious methodological weaknesses.

Table 1 presents a summary of the studies that have evaluated SIT. In the following discussion, these studies have been divided on the basis of (a) the nature of the subjects and (b) whether or not they meet the criterion proposed by Myers and Hammill (1982) that an effective evaluation should have an experimental-control group design with at least 10 experimental subjects.

RESEARCH ON SIT

Subjects with Learning Disabilities

Of the studies that involved subjects with LD and that met the Myers and Hammill (1982) criterion, those of White (1979) and De Pauw (1978) found positive results, while those of Ayres (1972b, 1978) and Jenkins and Sells (1984) found mixed results. Three studies that involved subjects with LD but that did not meet the Myers and Hammill criterion (Bullock & Watter, 1978; Schroeder, 1982; Ziviani, Poulsen, & O'Brien, 1982) found largely positive effects.

Subjects with Mental Retardation

Of the studies involving subjects with mental retardation (MR) that met the Myers and Hammill criterion, those of Morrison and Pothier (1972) and Montgomery and Richter (1977) found mixed results, while a second study by Morrison and Pothier (1978) found negative results. Two studies that did not meet the Myers and Hammill criterion, one involving children with MR and the other involving nondisabled children, found positive results (Culp, Packard, & Humphrey, 1980; Magrun, McCue, Ottenbacher, & Keefe, 1981).

The differences in the outcomes of these studies may be accounted for both by the methodological inadequacies of some of the studies and by significant differences in methodology. For example, only 3 of the 14 studies reviewed here included a control group for the effect of expectation and attention (Montgomery & Richter, 1977; Morrison & Pothier, 1972, 1978) and only 3 studies provided adequate reports of both perceptual-motor and academic variables (Jenkins & Sells, 1984; Schroeder, 1982; Ziviani et al., 1982). A significant proportion of the studies failed to assign subjects at random to control groups (e.g., Bullock & Watter, 1978), failed to control for therapist variables such as degree of therapist commitment to the program (e.g., Ayres, 1978), or had the same therapists involved in both the treatment and posttesting of the children (e.g., Montgomery & Richter, 1977). A number of studies had high subject mortality rates that they failed to explain (e.g., Bullock & Watter, 1978), provided inadequate controls for the effects of regression toward the mean, or provided inadequate information on program duration and subject selection (e.g., Ayres, 1972b).

In addition to the methodological problems listed above, there were significant variations among the studies in subject characteristics, variables used to evaluate the program, measures used to quantify these variables, program content, duration and hours of training, the ratio of therapists to children, and the length of

time between treatment and posttesting. To illustrate the significance of one of these differences, there is evidence to suggest that visual-motor variables are more likely to show positive results than sensory-motor and cognitive-language variables (Myers & Hammill, 1982).

The lack of consistency in the findings of the studies designed to evaluate SIT can be explained by variations in the methodology that occur from one study to another and by the design inadequacies of some of the studies. However, if the findings of the research on SIT are compared with the research on perceptual-motor therapies, a surprisingly consistent picture begins to emerge. The majority of studies that can be considered well designed according to the Myers and Hammill (1982) criterion show no significant differences between treatment and control groups, while the remainder are equally divided between mixed and positive results. Nuthall (1976) noted a similar, consistent pattern of findings in research evaluating a variety of different classroom teaching methods. He suggested that this pattern is a product of inadequate understanding and control of the processes involved in the methods being evaluated.

There is a lack of understanding of how variations in the processes of therapy

TABLE 1
A Summary of the Studies Evaluating Sensory Integration Therapy

Researcher and date	Population	Age	Subjects[b]	Hours of training	Comparison program	Dependent variables	Results
Ayers (1972b)	Learning disabled subgroup had auditory language problems	School age	42/84	UTE	No	Academic	+0
Ayres (1977)	Learning disabled with choreoathetoid movements	6 to 10	31/54	60[a]	No	Motor accuracy test of SCSIT	0
Ayres (1978)	Learning disabled subgroup had hypoactive nystagmus	6 to 10	46/92	60[a]	No	Academic	+ for group with hypoactive nystagmus; 0 for entire group
Bullock & Watter (1978)	Minimal brain damaged children	Preschool and school age	78/85	UTE	No	Clinical observations of neuro-development	+
Culp, Packard, & Humphrey (1980)	Nondisabled preschoolers	3 to 5	8/24	10	Yes—group received verbal body-part identification	Body concept development	+
De Pauw (1978)	Aphasics	Preschool	11/42	47[a]	Yes—remedial program	Perceptual-motor	+
Jenkins & Sells (1984)	Motor delayed—some mentally retarded	3 to 15	15/45	10 and 30	No	Gross motor, fine motor, academic	+0 +0 0
Magrun, McCue, Ottenbacher, & Keefe (1981)	Developmentally retarded	3 to 10	10/10	2	No	Frequency of verbalizations	+
Montgomery & Richter (1977)	Trainable mentally retarded	5 to 12	25/75	48	Yes—developmental PE program (small group)	Gross motor, fine motor, reflex integration	+ 0 +
Morrison & Pothier (1972)	Mentally retarded	Preschool	9/27	UTE	Yes—social reinforcement for casually selected gross motor activities	Gross motor, fine motor, language, personal-social	+ 0 + 0
Morrison & Pothier (1978)	Mentally retarded—12 with bilingual background	2.5 to 6	10/30	96	Yes—movement training group	Language, gross motor	0 0
Schroeder (1982)	Children with perceptual deficits	6	5/15	24	Yes—perceptual skills training	Auditory discrimination, academic, perceptual-motor	0 +0 +
White (1979)	Learning disabled	5 to 6	9/19	24	No	Reading	+
Ziviani, Poulsen, & O'Brien (1982)	Learning disabled boys	5 to 13	8/16	19.5	Remedial class activities	Gross motors fine motor, academic	+ + 0

UTE = unable to estimate; SCSIT = Southern California Sensory Integration Tests; PE = physical education; + = significant differences favoring the experimental group; 0 = no significant differences; +0 = mixed results.
[a]These figures are estimated. [b]This figure represents the number of experimental subjects compared to the total number of subjects.

might affect the results, or how specific therapies might have different effects on different children. Almost all research on SIT has focused on the apparently simple question of whether SIT is effective or not. Research has not sought to answer questions about how the components of the program or the context in which it is being used might influence children's learning during therapy.

It would seem that one critical factor is the nature of subjects. The ways in which children with LD differ from one another are numerous, and are likely to be related to their ability to gain from SIT. For example, subjects in different studies have been noted as having auditory language problems (Ayres, 1972b), reading failure (White, 1979), aphasia (De Pauw, 1978), emotional disturbance or communication disorders (Ayres, 1978; Jenkins & Sells, 1984), and English as a second language (Morrison & Pothier, 1978). Although Ayres's research has focused on children with LD, it is not clear exactly which children she identifies by this term.

Two studies have specifically addressed the question of which children are best suited to SIT therapy (Ayres, 1978; Jenkins & Sells, 1984). Ayres (1978) suggested that children with average or hyporeactive nystagmus benefit more than children with hyperreactive nystagmus. Jenkins and Sells (1984), on the other hand, suggested that SIT was best suited to children who exhibited moderate, but not mild, motor delay. Both of these studies, however, have serious methodological weaknesses that throw doubt on the validity of their findings (Densem, 1985).

The question then remains open as to whether SIT is an effective program for the treatment of children with a particular pattern of learning disabilities or related perceptual-motor difficulties.

The study reported here was an evaluation of the SIT program being run by the Occupational Therapy Department of the Christchurch Hospital, which took as one of its major purposes the identification of those children who seemed most likely to benefit from such a program. The study was initiated in response to a request from the hospital authorities to evaluate the program, which had been running for a number of years and had accumulated a considerable waiting list of children referred for therapy.

METHOD

Subjects

The subjects were 57 5- to 10-year-old children who had been referred to the the ocupational therapists' sensory integrative program at Christchurch Hospital between August 1982 and February 1984. These constituted the entire population of children between these ages referred to the program, excluding those early referrals who had been given an urgent rating, and who had already been treated, and those children (eight) whose parents decided that their children no longer required therapy. A further seven children were excluded by the principal researcher and the therapists: three because they appeared to have no sensory integrative difficulties following testing on the Southern California Sensory Integration Tests (SCSIT), two because of their concurrent involvement in other intervention programs, one because her motor problem was due to physical handicap, and one because of mental retardation and difficulties with testing. The remaining subjects exhibited a wide array of handicapping conditions, including mild mental retardation, behavioral disturbance, mild cerebral palsy, and epilepsy.

Children were referred to the SIT program predominantly because of perceptual-motor difficulties (84%) and learning disabilities (33%). A very large proportion of the children, however, had difficulties with both school achievement and with motor activities. Fifty-four of the fifty-seven children were described by their teachers as having difficulties with school achievement. The same number of children were described by either their teachers or their parents as being uncoordinated or clumsy. A large proportion of the children (89%) were reported to have evidence of social or behavioral problems either at school or at home, and a similarly high proportion (89%) were described as having difficulties with concentration. According to their teachers, about half the children had problems with copying from the blackboard. Speech problems were reported in 85% of the children, and 60% had attended speech therapy. However, according to their teachers, 78% of the children were estimated to be average or above average in intelligence.

These 57 children were randomly assigned to three groups of 19 children each. These groups were (a) an occupational therapists' sensory integrative treatment group (OT), (b) a physical education control group (PE), and (c) a no-treatment control group (NT). Subsequent to group assignment, two children from the physical education control group were withdrawn from the program by their parents.

So that equal numbers of older and younger children of each sex were assigned to all groups, subjects of each sex were ordered according to their age and then divided into two groups at the median age (88 months). Subjects within each age and sex category were then randomly assigned to the three treatment groups (OT, PE, and NT).

Table 2 presents the mean age, number of program sessions attended, socioeconomic status levels, sex, and the number attending special class or having other treatment, at the time of posttesting.

Procedure

Pretesting was conducted in February 1984 by the two pediatric occupational therapists at Christchurch Hospital and by an assistant clinical psychologist, who also had experience as a classroom teacher. Each subject had three assessment sessions: two with an occupational therapist and one with the assistant clinical psychologist.

The OT and PE programs began in March and continued until the second week in July, excluding a 2-week May vacation. The mean number of training sessions, each an hour long, was 17.4 for the children in the OT group and 16.0 for the children in the PE group.

Posttesting began in October 1984, following a 12-week latency period, and lasted 4 weeks. The psychologist and two occupational therapists, who had had no contact with the experimental program, conducted posttesting. Examiners at both pre- and posttesting were kept unaware of the experimental or control group status of the subjects. Particular care was

taken at posttesting so that the children did not inadvertently reveal their group membership, and in all except a few cases, this was successful. Following posttesting, a partial crossover was carried out in which the children in the PE and in the no-treatment groups received the OT treatment, while those already treated with SIT went on to a no-treatment condition.

Measurement of Dependent Variables

The following instruments were individually administered to each subject as pre- and posttest measures.

Language. The Bankson Language Screening Test (BLST) (Bankson, 1977) consists of a battery of 17 nine-item subtests organized into five general categories: semantic knowledge, morphological rules, syntactic rules, visual perception, and auditory perception. The test was developed to measure expressive language and the psycholinguistic and perceptual skills on which this is based. The test is suitable for children from 4 to 8 years, but is more sensitive to developmental differences at the younger ages (Bankson, 1977). Test reliability is reported to be between .94 and .96. Correlations between BLST and other widely used language tests range from .54 to .64 (Bankson, 1977).

Perceptual-Motor and Sensory Integrative Skills. The battery of Southern California Sensory Integration Tests (SCSIT) (Ayres, 1972c, 1980) was used as a pre- and posttest measure and, for the children in the OT groups, to establish the patterns of deficits on which their individual intervention programs were based. SCSIT is designed to help diagnose children's dysfunctions in the areas of form and space perception, tactile perception, postural and bilateral integration, and motor skills.

These tests, however, have a number of limitations. First, the normative data are of questionable value (Westman, 1978) and may not be applicable to New Zealand populations (Ritchie, personal communication, 1985). In the present study, SCSIT raw scores, rather than standardized scores, have therefore been reported. The test-retest reliabilities range from .01 to .89 (Ayres, 1980), and it has been suggested that the standard error of measurement and the reliability of each test be checked for age-related versus random variability, before crediting a score with significance (Ayres, 1980; Westman, 1978). Third, as Ayres (1980) notes, some of the tests are limited by ceiling and floor effects. The motor accuracy (MAC) and design copying (DC) subtests are the most reliable, but DC can be affected by previous practice (Ayres, 1978; Westman, 1978). The validity of these tests and their related syndromes has, furthermore, not been well established (Bochner, 1980; Reed, 1978; Westman, 1978). The SCSIT scores were included, in spite of their low reliability and questionable validity, because they were the tests used by the occupational therapists at Christchurch Hospital for diagnosis, treatment planning, and assessment of progress in treatment.

Since averaging the group means for the MAC-R (right) and MAC-L (left) subtests mixes together the scores from subjects' dominant and nondominant hands, the motor accuracy test scores were classified by dominant hand (MAC-Dom) and by nondominant hand (MAC-NDom) rather than by left and right hands.

Reading. Each child was asked to read a sample of prose extracts that were below, at, and above their estimated reading age, and their reading miscues were recorded. The reading was tape-recorded for later reliability checks. Reading ages for those capable of reading were computed by establishing the difficulty level of the extract at which the children read with 95% accuracy (instructional level, Clay, 1979). This procedure provided a detailed record of each child's reading behavior and also had the advantage of being applicable to a broad range of reading abilities. However, because the step between reading at the emergent level and reading at the beginner reader level is quite different from other intervals (for example, the interval between reading 5½- to 6-year material and 6- to 6½-year material; see Note 1), the progress of readers who were reading at the emergent level at the time of pretesting was analyzed both separately from and together with the progress of other readers. To check the reliability of the reading age measures, one-quarter of the subjects were randomly selected and their miscues recoded from the taped reading passages by a second rater. The interrater reliability for number of miscues, calculated by dividing the number of agreements by the total number of agreements and disagreements, was 82%.

For children who were between the ages of 5 and 7 years, and who were reading between the emergent and 5½-

TABLE 2
Characteristics of Experimental and Control Groups Prior to Posttesting

	Group		
	NT	OT	PE
n	19	19	17
Female *n*	3	4	3
Special Class or other treatment	0	3	0
Age[a]			
Mean	7.10	8.1	7.10
SD	1.6	1.3	1.4
Sessions			
Mean	0	17.42	16.0
SD	–	0.50	1.8
SES[b]			
Mean	3.5	3.6	3.5
SD	1.6	1.4	1.1

[a]Age as of February 1, 1984. [b]SES = Socioeconomic status; based on Elley and Irving (1976) Index, a 7-point scale. NT = no treatment; OT = occupational therapy; PE = physical education.

year level, a letter identification and a word test (Clay, 1979) were also administered. This gave a more accurate and sensitive measure of these children's reading development. In the letter identification test, children were asked to identify the lowercase and uppercase letters of the alphabet. The word test is a list of 15 words compiled from the 45 most frequently occurring words in a common early reading series. This has been found to be a good instrument for estimating the progress of children during the first year of instruction and for retarded readers in the second year (Clay, 1966).

Handwriting. The instrument used to measure difficulties with the fine motor coordination involved in handwriting was a specially constructed coding schedule, based on the Slingerland (Slingerland, 1970) and the Purdue Perceptual-Motor Survey (Roach & Kephart, 1966). Children wrote their names and addresses and then copied either one of two writing samples (depending on age level) within a 10-minute time limit. The number of letters or letter-like shapes were then totaled, and the errors were coded according to 17 categories that were grouped according to whether they were perceptual errors (e.g., reversals, inversions), letter misformations (e.g., incompleteness of letters, choreoathetosis), or spacing or word order errors. Two examiners were trained in the use of the coding schedule, and they separately coded all writing samples. Where there was disagreement, an average of the two scores was taken as the final score. The interrater agreement for the total number of errors was 88.3% and ranged from 83.9% to 86.2% in the three categories. Subcategory agreement was from 42.9% to 91.7%. The number of errors was divided by the total number of letters recorded, to adjust for speed of writing. This is referred to as the "writing error ratio." Correlations between the writing error ratio and Southern California Sensory Integration Motor subtests ranged from .53 to .69.

Self-Esteem. The Culture-Free Self-Esteem Inventories for Children, Form A (Battle, 1981) is a self-report scale developed to measure an individual's perception of self and includes 50 items relating to school, peer, parent, and general self-esteem issues and 10 lie items.

Unfortunately, the pretest–posttest correlations of scores on the self-esteem inventory were extremely low, which indicated that the reliability of the test with this sample was lower than reported by Battle (1981); this made meaningful interpretation of results impossible.

Behavior, Social Adjustment, and School Progress. Parent and teacher perceptions of the children's behavior, social adjustment, and school progress were obtained via two separate questionnaires. Individual interviews were carried out with the parent (or parents) at the time they filled in the questionnaires, and information about each child's previous or present involvement in additional treatments, such as remedial reading, was obtained. A sample of this information was checked for accuracy against school records. Both parents and teachers were asked to indicate on a 4-point scale the frequency of a number of social and classroom behaviors. Teachers were also asked to rate the child's abilities at reading, spelling, handwriting, mathematics, physical education, writing, language, and speech on a 5-point scale.

Intervention Programs. Procedures developed predominantly by Ayres (1972a), but also by Bobath and Bobath (1976), Kephart (1960), and Rood (1962), were followed with the subjects of the OT group. The program involved activities for improving children's balance, fine and gross motor skills, responsiveness to touch, integration of both sides of the body, and motor movement planning. These activities included riding a scooter-board down a ramp, games involving fine motor skills (such as puzzles), obstacle courses, hitting a swinging ball, and balance-beam work. The program was carried out with each child individually.

The PE program involved activities from the New Zealand Department of Education (1978) publications on physical education for schools. Each session involved fitness activities, the teaching and practice of physical skills, and games. Skill units that were taught included small and large ball handling, summer and winter game skills (older group), and folk dancing (two younger groups). The fitness component involved running and skipping and jumping activities, followed by fitness exercises, such as bend-claps (an exercise involving bending the knees and at the same time clapping the hands with the arms stretched high over the head). The children met in small groups of five or six children for the PE program.

The children in the no-treatment control group (NT) had no contact with the OT program between pretesting in February and posttesting in October. The children did, in the natural course of their schooling, come in contact with the educational remedial systems. Information about these contacts was obtained at the time of posttesting.

RESULTS

The significance of the differences between the three groups on the pretest scores was analyzed using analysis of covariance with age as the covariate (Keppel, 1982). Significant differences were found on two of the SCSIT subscores: Bilateral Motor Coordination, $F(2, 48) = 7.2, p < .005$; Motor Accuracy Dominance Test, $F(2, 53) = 4.15, p < .05$; and on the reading word test, $F(2, 25) = 8.10, p < .01$. In each case, the OT group had the lowest scores. These differences were taken into account by including the pretest scores as covariates in the analysis of the posttest scores.

Reliability

The reliability of all measures, except those for which it was possible to assess rater reliability independently (e.g., the handwriting error ratio), was estimated from the correlation between pretest and posttest scores.

Correlations of 0.5 and above were taken as an indication that a minimum level of reliability had been achieved. Because there are special difficulties involved in using many of the instruments with children with sensory motor problems, scattergrams of the distributions of pretest and posttest scores for each group on all measures were carefully examined for floor and ceiling effects and for unusual individual score patterns. This

provided a check on the consistency of the scores between groups, and the consistency of individual children's scores between pre- and posttest. Only those measures that achieved a minimum level of reliability and consistency using the above criteria have been reported below.

Examination of scattergrams revealed that the scores of 3 subjects (1 from each group) showed exceptional variations from pretest to posttest, suggesting that there were problems in the reliability of the scores of these children. For example, on the Double Tactile Stimulation subtest of the SCSIT, the score of one child, who had recently developed epilepsy, fell more than 2 standard deviations. The scores of these 3 subjects were excluded from the analysis. In addition, at the time of posttesting, it was found that 1 child from the PE group had moved to another city.

These omissions left a total of 54 children; 19 in the OT group, 16 in the PE group, and 19 in the no-treatment control group.

Comparisons Between Groups at Posttesting

Table 3 presents unadjusted means and standard deviations for pre- and posttest scores for each group on the measures that showed adequate levels of reliability. To determine the significance of the differences between groups, the posttest scores of the three groups on each of the measures were subjected to analysis of covariance, with pretest scores and age as covariates. Table 4 contains the posttest means for each group adjusted for pretest and age, and the *F*-ratio and significance level of the between group differences.

No significant differences between groups were found on the Southern California Sensory Integrative Tests, the Bankson Language Screening Test, $F(2, 49) = 2.59$, $p < .10$ and $p > .05$, and the handwriting error ratio, $F(2, 47) = 0.38$, $p > .05$.

No significant effect of treatment was shown in reading age scores when all readers were considered together, $F(2, 47) = 2.27$, $p > .05$. However, the results of the analysis of the reading age scores for subjects already reading at pretest demonstrated a significant effect of treatment group, $F(2, 24) = 5.37$, $p < .01$. A comparison between the OT

TABLE 3
Means and Standard Deviations of Pre- and Posttest Scores

Variable		Group					
		NT		OT		PE	
		$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*
Handwriting[a]	Pre	1.00	0.66	0.92	0.58	0.96	0.57
	Post	0.65	0.45	0.76	0.52	0.65	0.62
Reading age	Pre	6.29	2.18	6.06	1.73	5.95	1.43
	Post	6.87	2.58	6.67	2.38	6.39	1.67
Reading word[b]	Pre	3.70	4.11	3.60	3.53	7.42	6.08
	Post	8.70	5.08	9.22	4.49	8.75	6.16
Language	Pre	107.74	29.28	107.42	32.46	106.58	33.00
	Post	116.26	23.27	114.47	28.09	119.56	24.61
SCSIT							
Position in space	Pre	14.05	5.45	12.47	5.86	13.88	4.00
	Post	16.05	4.34	15.68	5.12	16.94	1.62
Design copying	Pre	8.84	12.01	8.16	6.35	9.59	5.98
	Post	9.84	4.56	9.26	6.21	10.50	5.90
Graphesthesia	Pre	8.21	3.88	8.89	4.48	7.77	4.38
	Post	10.15	3.70	9.79	5.40	7.31	4.92
Limitation of postures	Pre	9.89	4.24	10.17	4.89	10.94	5.61
	Post	11.21	4.09	11.50	5.52	11.50	4.20
MACR	Pre	140.68	12.46	133.68	13.32	139.18	12.22
	Post	143.56	10.12	139.00	12.37	140.25	12.43
MACL	Pre	133.63	14.00	134.95	11.20	139.17	10.05
	Post	135.89	11.05	134.68	10.98	137.31	10.16
MAC-Dom	Pre	141.37	12.43	136.05	12.34	143.29	10.11
	Post	144.44	9.15	140.16	11.79	140.56	12.83
MAC-NDom	Pre	132.95	13.63	132.58	12.05	135.06	10.59
	Post	135.00	11.17	133.53	11.01	137.00	9.54

Only those measures reaching reliability criteria are reported. NT = no treatment; OT = occupational therapy; PE = physical education; SCSIT = Southern California Sensory Integrative Tests; MAC = Motor Accuracy; L = left hand, R = right hand; Dom = dominant hand; NDom = nondominant hand. [a]On these measures, zero is the highest possible score. [b]Beginner readers only.

TABLE 4
Adjusted Posttest Means and Significance of Between-Group Differences

	Group				
	NT	OT	PE	*F* ratio	*df*
Handwriting[a]	.64	.73	.68	0.379	2/47
Reading age	6.51	6.93	6.21	2.22	2/47
Reading age[b]	7.63	8.37	8.03	5.37*	2/24
Reading word test[c]	8.36	11.85	6.62	2.316	2/22
Language	115.52	115.19	121.05	2.590	2/49
SCSIT					
Position in space	16.12	16.36	16.74	0.184	2/48
Design copying	10.41	9.69	10.15	0.348	2/48
Graphesthesia	10.37	9.48	8.22	1.501	2/49
Imitation of postures	11.56	11.50	11.63	0.006	2/49
Motor accuracy-R	142.00	143.24	139.39	1.360	2/48
Motor accuracy-L	136.94	136.26	135.88	1.010	2/48
Motor accuracy-Dom	143.91	144.01	139.10	2.836	2/48
Motor accuracy-NDom	135.18	135.40	136.30	6.167	2/48

Means adjusted for pretest score and age. NT = no treatment; OT = occupational therapy; PE = physical education; SCSIT = Southern California Sensory Integrative Tests; L = left hand; R = right hand; Dom = dominant hand; NDom = nondominant hand, MAC = motor accuracy. [a]On these measures, zero is the highest possible score. [b]Ss reading beyond the emergent level at pretest only. [c]Emergent readers only. *p = .01.

and NT groups, using the Scheffé test, revealed a significant difference, $F(2, 24) = 6.64, p < .01$. A comparison between the OT and PE groups also showed a significant difference, $F(2, 24) = 2.12, p < .05$. The OT group had significantly greater adjusted posttest means than both the PE and the NT groups. The difference between the PE and NT groups, $F(2, 24) = 3.65, p < .01$, was also significant and favored the PE group.

It is clear from the results of these analyses that, with the exception of the reading scores for those children who were already reading at pretesting, none of the differences between the three groups was statistically significant. This means that, for most of the measures when children's age and status at the time of pretest are taken into account, the progress of children in the different treatment programs did not differ significantly.

Identifying Children Likely to Benefit from SIT

The second major purpose of the study was to identify those children most likely to benefit from the SIT program. It was possible that, because of the wide range of handicapping conditions evident among the children in the program, the average effect was obscuring the fact that some children were making significant progress. Two post hoc analyses of the data were undertaken to try to identify whether there existed, within the OT group, a clearly distinguishable subgroup of children who had made substantial gains.

The first analysis involved the use of multiple regression analyses, with each reliable pretest measure used as a potentially significant independent variable. The one variable that was found to be most frequently related to the progress of the OT group children on the dependent measures was the SCSIT subtest, Double Tactile Stimulation (DTS). Other variables that were related to children's progress on some of the dependent variables were evidence of severe epileptic seizures, the sensory integrative syndrome classification developed by the occupational therapists, ability to copy from the blackboard in school, scores on MAC-L subtest of the SCSIT, and a group of measures reflecting the extent of sociocultural deprivation. Evidence of epilepsy was highly correlated with the DTS scores.

The second approach to the problem of identifying the distinguishing characteristics of children not likely to make progress in the OT program was to contrast those who made the most progress with those who made none at all. This was done by calculating the residual scores obtained from the multiple regression analyses of the criterion scores. An examination of these residual scores indicated that 6 children had standardized residual scores of 1 standard deviation or more below the expected mean on at least 3 of the 14 reliable criterion measures.

An examination of the pretest scores of these 6 subjects showed that they (a) had very low scores on the DTS, (b) were epileptic, (c) had a mother who was a solo parent on a single-parent benefit, or (d) had severe behavior problems as rated by the occupational therapists.

The 3 children who were epileptic all had low DTS scores at pretesting. The criterion of epilepsy therefore replicated that of low DTS scores, except in one case, in which the child's epilepsy did not become evident until after pretesting.

DISCUSSION

The results of this study show that the sensory integrative therapy program implemented by the Occupational Therapy Department of Christchurch Hospital produced no greater gains in language development, in all of the areas of perceptual-motor development that could be reliably measured, or in handwriting skills, than did a parallel physical education program or no treatment at all. The only area in which the program showed an effect was in the reading progress of children who were already reading at the time of pretest. There was, however, no effect on the reading progress of those children who could not yet read at the time of the pretest.

These findings are congruent with the findings of earlier research. Few adequately designed studies have found a consistent pattern of results favoring a sensory integrative therapy program.

Moreover, when there are inconsistent findings, those measures least likely to show any effect are the sensory-motor test results.

It should be noted, however, that there are considerable difficulties in measuring motor skills accurately (Bochner, 1980; Jenkins & Sells, 1984). The majority of the SCSIT subtests are of questionable reliability and validity. There are no reliable measures of somatosensory perception and only one reliable measure of gross motor development (Imitation of Postures).

In trying to identify those children who were most likely to benefit from the sensory integrative therapy program, one factor (the measure of Double Tactile Stimulation) did appear to be consistently related to most of the criterion measures. Further analysis suggested that those children who had epilepsy, significant behavioral problems, or whose parent was on a single-parent government benefit were unlikely to show progress during therapy.

It has been suggested that a low DTS score in children over 6 years may be evidence of central nervous system pathology (Ayres, 1972c; Fink & Bender, 1953) and low mental age (Swanson, 1957). It is probable that those children whose parent was on a single-parent government benefit suffered from a range of social deficits usually associated with low income and difficult living conditions. Children with significant behavioral problems probably spent significantly less time actually engaged in the planned components of the sensory integrative therapy program.

What this analysis suggests is that those children who showed little or no progress at all during the sensory integrative therapy program were those who suffered from a range of handicapping conditions, each requiring a different kind of intervention. It does not, however, throw much light on how the SIT program might be working for those children who do seem to have made substantial progress.

This highlights the major limitation of a comparative evaluation study of this type. A general answer to the question of whether SIT is effective or not can never be found. Every evaluative study based on outcome measures is dependent on the specific combination of subjects, program content, dependent measures, and other contextual factors that are involved in the particular study. A number of authors (e.g., Barbatsis, 1978; Tuckman, 1978) have suggested that comparative effectiveness studies are premature in areas of inquiry in which the significant variables and models of the process of change have not been clearly identified. The accumulation of further comparative effectiveness studies tends to produce a more or less random distribution of outcomes (Nuthall, 1976) and little progress in understanding the process of therapy. The research focus, therefore, needs to shift toward relating specific treatment variables to the process of development and learning during treatment, as well as to outcome measures. The question then becomes, not "How effective was the program?" but rather, "How does it work and for whom?"

ABOUT THE AUTHORS

Julie Densem, *MA, is a psychologist, Psychological Service of the New Zealand Justice Department. This study is based on her master's thesis.* **Graham Nuthall,** *PhD, is professor of education at the University of Canterbury.* **John Bushnell,** *MA, is a psychologist, Department of Community Health, University of Otago, Christchurch School of Medicine.* **Jacqueline Horn,** *PhD, is a psychologist at Sunnyside Hospital, Christchurch. Address: Julie Densem, Psychological Service, PO Box 13-223, Armagh, Christchurch, New Zealand.*

ACKNOWLEDGMENTS

1. This research was supported in part by the J.R. McKenzie Foundation, the Philip Brown Trust, The Todd Foundation, The Sargood Bequest, and the Occupational Therapy Association.

2. We gratefully acknowledge the assistance of the three pediatric occupational therapists at Christchurch Hospital: Mrs. Marie Chester, Mrs. Sally Wilson, and Ms. Elizabeth Saunders, for data collection and implementation of the intervention program; Mrs. Helen Leahy for the planning and execution of the physical education program; the contact organization for their practical help; and Mr. Neville Blampied for useful comments on the original draft.

NOTE

1. Children begin reading in New Zealand schools on entry at 5 years of age.

REFERENCES

Ayres, A.J. (1970). Interpretation of Southern California Sensory Integrative Tests. Los Angeles: Western Psychological Services.

Ayres, A.J. (1972a). Sensory integration and learning disorders. Los Angeles: Western Psychological Services.

Ayres, A.J. (1972b). Improving academic scores through sensory integration. Journal of Learning Disabilities, 5, 338–344.

Ayres, A.J. (1972c). Southern California sensory integrative tests. Los Angeles: Western Psychological Services.

Ayres, A.J. (1977). Effect of sensory integrative therapy on the co-ordination of children with choreoathoid movements. The American Journal of Occupational Therapy, 31, 291–293.

Ayres, A.J. (1978). Learning disabilities and the vestibular system. Journal of Learning Disabilities, 11, 18–29.

Ayres, A.J. (1980). Southern California sensory integrative tests manual (Rev. ed.). Los Angeles: Western Psychological Services.

Bankson, N. (1977). Bankson language screening test. Austin, TX: PRO-ED.

Barbatsis, G. (1978). The nature of inquiry and analysis of theoretical progress in instructional television from 1950–1970. Review of Educational Research, 48, 399–414.

Barsch, R. (1967). Achieving perceptual-motor efficiency: A space-oriented approach to learning. Seattle: Special Child Publications.

Battle, J. (1981). Culture-free self-esteem inventories for children and adults. Seattle: Special Child Publications.

Bobath, K., & Bobath, R. (1976). Cerebral palsy. In P. Pearson & C. Williams (Eds.), Physical therapy services in the developmental disabilities. Springfield, IL: Thomas.

Bochner, S. (1980). Sensory integration therapy and learning disabilities: A critique. Australian Occupational Therapy Journal, 2, 125–138.

Bullock, M.J., & Watter, P. (1978). A study in the effectiveness of physiotherapy in the management of young children with minimal cerebral dysfunction. Australian Journal of Physiotherapy, 24, 111–119.

Clay, M.M. (1966). Emergent reading behavior. Unpublished doctoral dissertation, University of Auckland.

Clay, M.M. (1979). The early detection of reading difficulties: A diagnostic survey with recovery procedures. Auckland: Heinemann.

Cratty, B.J., & Marlin, M. (1969). Perceptual-motor efficacy in children. Philadelphia: Lea & Febiger.

Culp, R., Packard, V., & Humphrey, R. (1980). Sensorimotor versus cognitive perceptual training effects on the body concept of preschoolers. The American Journal of Occupational Therapy, 34, 259–262.

Densem, J.F. (1985). An evaluation of the Christchurch Hospital's occupational therapy programme for children with perceptual-motor defects. Unpublished master's thesis, University of Canterbury, Christchurch, New Zealand.

De Pauw, K.P. (1978). Enhancing the sensory integration of aphasic students. Journal of Learning Disabilities, 11, 142–146.

Doman, R., Spitz, E., Zucman, E., Delacato, C., & Doman, G. (1960). Children with severe brain injuries: Neurological organizations in terms of

mobility. Journal of the American Medical Association, 174, 257–262.

Elley, W.B., & Irving, J.C. (1976). Revised socio-economic index for New Zealand. New Zealand Journal of Educational Studies, 12, 152–160.

Fink, M., & Bender, M. (1953). Perception of simultaneous tactile stimuli in normal children. Neurology, 3, 27–34.

Frostig, M. (1967). Education of children with learning disabilities. In E.C. Prierson & W. Barbe (Eds.), Educating children with learning disabilities. New York: Appleton-Century-Crofts.

Getman, G., Kane, E., Halgren, C., & McKeen, G. (1968). Developing learning readiness. Manchester, MO: McGraw-Hill.

Goodman, L., & Hammill, D. (1973). The effectiveness of the Kephart-Getman activities in developing perceptual-motor and cognitive skills. Focus on Exceptional Children, 4, 1–9.

Green, C., Hope, M., Oates, R.K., Parry, T., & Procopis, P. (1982). Report to the Australian College of Paediatrics on Sensory Interactive Therapy.

Jenkins, J., & Sells, C. (1984). Physical and occupational therapy: Effects related to treatment frequency and motor delay. Journal of Learning Disabilities, 17, 89–95.

Kephart, N. (1960). The slow learner in the classroom. Columbus, OH: Merrill.

Keppel, G. (1982). Design and analysis: A researcher's handbook (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Magrun, W.M., McCue, S., Ottenbacher, K., & Keefe, R. (1981). Effects of vestibular stimulation on spontaneous use of verbal language in developmentally delayed children. The American Journal of Occupational Therapy, 35, 101–104.

Montgomery, P., & Richter, E. (1977). Effect of sensory integrative therapy on the neuromotor development of retarded children. Physical Therapy, 57, 799–806.

Morrison, D., & Pothier, P. (1972). Two different remedial motor training programs and the development of mentally retarded preschoolers. American Journal of Mental Deficiency, 77, 251–258.

Morrison, D., & Pothier, P. (1978). Effects of sensory-motor training on the language development of retarded preschoolers. American Journal of Orthopsychiatry, 48, 310–319.

Myers, P., & Hammill, D. (1982). Methods for learning disorders. New York: John Wiley.

New Zealand Department of Education. (1978). Physical Education–Standard 2 to Form 2. Wellington, New Zealand: Government Printer.

Nuthall, G.A. (1976). Which methods of teaching should a teacher use? Unpublished manuscript, University of Canterbury, Christchurch, New Zealand.

Ottenbacher, K. (1982). Sensory integration therapy: Affect or effect? The American Journal of Occupational Therapy, 36, 571–578.

Reed, B.C. (1978). Review of the Southern California Sensory Integration Tests. In O.K. Buros (Ed.), The eighth mental measurements yearbook (pp. 1407–1408). Highland Park, NJ: Gryphon Press.

Roach, E., & Kephart, N. (1966). The Purdue perceptual-motor survey. Columbus, OH: Merrill.

Rood, M. (1962). The use of sensory receptors to activate, facilitate and inhibit motor response, autronomic and somatic in developmental sequence. In C. Sattley (Ed.), Approaches to the treatment of patients with neuromuscular dysfunction. Dubuque, IA: Brown.

Schroeder, R. (1982). Improvement in academic achievement through enhancement of perceptual and sensory integrative functioning. School Psychology International, 3, 97–104.

Sieben, R.L. (1977). Controversial medical treatments of learning disabilities. Academic Therapy, 13, 133–147.

Slingerland, B.M. (1970). Slingerland screening tests for identifying children with specific language disability. Cambridge, MA: Education Publishing Service.

Swanson, R. (1957). Perception of simultaneous tactile stimulation in defective and normal children. American Journal of Mental Deficiency, 61, 743–752.

Tuckman, B.W. (1978). Conducting educational research (2nd ed.). New York: Harcourt Brace Jovanovich.

Westman, A.S. (1978). Review of Southern California Sensory Integrative Tests. In O.K. Buros (Ed.), The eighth mental measurements yearbook (pp. 1408–1409). Highland Park, NJ: Gryphon Press.

White, M. (1979). A first grade intervention program for children at risk for reading failure. Journal of Learning Disabilities, 12, 231–237.

Ziviani, J., Poulsen, A., & O'Brien, P. (1982). Effect of a sensory integrative/neurodevelopmental programme on motor and academic performance of children with learning disabilities. Australian Occupational Therapy Journal, 29, 27–33.