Reexamination of Sensory Integration Treatment: A Combination of Two Efficacy Studies

Bonnie J. Kaplan, Helene J. Polatajko, Brenda N. Wilson, and Peter D. Faris

Little empirical support exists for the application of sensory integration treatment (SIT) to assist children with learning problems. Treatment efficacy studies are expensive and difficult to carry out, and they have necessarily employed small samples that are inevitably heterogeneous. We have reanalyzed the efficacy of SIT by combining the data from one study involving 29 children in Alberta and a second study involving 67 children in Ontario. The results from each individual study, and now the results from the combined study, lead one to the conclusion that the therapeutic effect of SIT on children with learning deficits is not greater than other, more traditional methods of intervention.

Ever since Ayres first began publishing her theory on sensory integration treatment (SIT) for children with learning problems, its value has been debated. Early reports of its efficacy (e.g., Ayres, 1972) were based on poorly controlled studies. When Ottenbacher (1982) reviewed the first decade of SIT experimental research, however, he concluded that there was sufficient evidence of its value to warrant its continued application to the problem of learning disabilities. Ottenbacher's recommendation was that further research be carried out, to provide the data necessary for firm conclusions. Two recent reviews have summarized the literature regarding (a) diagnosis of sensory integrative dysfunction and (b) treatment efficacy. In the first category, Cummins (1991) reviewed Ayres's factor analytic studies that provided the framework for her ideas regarding diagnosis and concluded that there was no established validity for the diagnostic procedures proposed by Ayres. In the second category, we recently evaluated the second decade of research on treatment efficacy and came to the conclusion that SIT had still not fulfilled its promise (Polatajko, Kaplan, & Wilson, in press). However, as we indicated in that review, very few well-controlled studies exist in the published literature, in spite of the large interest in the topic.

One reason for the scarcity of studies is simply that treatment efficacy studies are difficult and expensive to carry out. One controlled evaluation study of SIT that included a comparison treatment (tutoring), with 29 children receiving treatment for approximately 75 sessions of one-to-one therapy, cost over $200,000 (Wilson, Kaplan, Fellowes, Gruchy, & Faris, 1992). Even when an expensive study of this type has been carried out, the answers that it can provide may be somewhat limited. Sample sizes are necessarily small. Because of the nature of children with learning or behavioral deficits, subject variability is high. Controlling for normal maturational changes is almost impossible because of the difficulty in recruiting for a no-treatment control group (Polatajko, Law, Miller, Schaffer, & Macnab, 1991).

Given these important issues (the controversial nature of SIT, the costliness of doing SIT research, and the difficulties in drawing definitive conclusions from any single, small sample study), we thought it might be scientifically useful to combine the data from two recently published and well-controlled studies to see if any further light could be shed on an important question: Can SIT be a useful treatment for children with learning problems?

Method

Description of the Studies

The first study providing data for these analyses was the Alberta one mentioned above (Wilson et al., 1992), in which 29 children were randomly assigned to either 75 sessions of SIT by qualified occupational therapists or 75 sessions of tutoring by special education teachers. The children were aged 5 years 2 months to 8 years 6 months (M = 6 years 8 months; SD = 12.6 months). Assessment of academic, motor, and sensory-perceptual skills occurred prior to treatment, halfway through (at about 6 months), and at the end of treatment (at 12 months).

The second study providing data for these analyses (the Ontario study) consisted of 35 children randomly assigned to receive SIT and 32 children randomly assigned to receive perceptual-motor therapy (PM) (Polatajko et al., 1991). At intake the children were between 6 years and 8 years 11 months. Treatment for each child consisted of 1 hour of therapy per week for 6 months, followed by 3 months with no therapy. Assessment of academic function, motor skills, and self-esteem occurred prior to treatment, at the end of treatment (at 6 months), and after a 3-month break (at 9 months).

Polatajko et al. (1991) had an additional 13 children randomly assigned to a no-treatment control group. Assignment to this group had to be terminated prematurely because it was preventing parents from agreeing to participate in the study. Hence, the data from this small control group have not been published and will not be included in the analyses in this presentation.

Both groups of investigators scrupulously attended to design features such as random allocation of subjects (with defined constraints), assessment of dependent measures by individuals blind to group assignment, and well-defined therapy manuals so that the treatments being offered did not overlap in content.

Subjects

Children referred to both studies met similar inclusion and exclusion criteria. Both samples would best be categorized in the category of developmental coordination disorder (American Psychiatric Association, 1987); that is, the children had coordination and clumsiness problems that interfered with their performance in academics and activities of daily living. All had been referred to physical or occupational therapy for this problem and, upon further occupational therapy assessment, were given a diagnosis of sensory integrative dysfunction. They were all of average intelligence, defined as a score of 85 or higher on the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974) in the Alberta study and as a score of 80 or higher on the Slosson Intelligence Test (Slosson, 1963) in the Ontario study. The academic performance of this group of children was delayed by 6 to 12 months on the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977) (Polatajko et al., 1991), or was below average on at least one of its three subtests (Wilson et al., 1992), which included the Reading Cluster, the Broad Cognitive Cluster, and the Perceptual Speed Cluster. Schaffer, Law, Polatajko, and Miller (1989) examined the performance of the Ontario sample and found that 80% and 91% of the subjects were delayed in reading and written language, respectively.

Exclusion criteria for both studies included having medical problems that interfered with treatment, or neurological disorders, such as epilepsy or cerebral palsy. In the Alberta study, children with attention deficit-hyperactivity disorder (American Psychiatric Association, 1987) were also excluded. More detailed descriptions of these subjects are available in Polatajko et al. (1991) and Wilson et al. (1992). However, it is clear that the subjects in Grades 2 and over were experiencing learning difficulties while the younger subjects had many characteristics of learning disabilities and were considered to be at risk for later learning problems.

Dependent Measures

The two studies being combined for analyses in this article were not designed with such a combination in mind. Hence, some of the variables that were assessed employed different measures and cannot properly be compared post hoc. On the other hand, the measures that can be combined cover a variety of academic and motor skills.

Because children in the Polatajko et al. (1991) study received only 6 months of treatment, the data that were used to determine treatment effects were the posttest data from the Ontario study and the midtest data from the Alberta study. The measure used to evaluate fine and gross motor skills in both studies was the Bruininks-Oseretsky Test of Motor Proficiency (Bruininks, 1978). From this, we employed two variables: fine motor composite and gross motor composite. The measure used to evaluate academic and visual motor skills in both studies was the Woodcock-Johnson Psycho-Educational Battery. From this test, we used the Reading Cluster, the Math Cluster, and the Preschool Cognitive Cluster to measure academic skills, and the Perceptual Speed Cluster to measure visual motor skills.

Results

Comparison of Samples at Pretest

Children who received SIT ($n=49$) from both the Alberta study and the Ontario study were compared with the tutoring group ($n=15$) from Alberta and the perceptual-motor group ($n=31$) from Ontario on a number of variables evaluated prior to any intervention. No difference was found in the distribution of males and females, $\chi^2(2, N=95)=0.28$, *ns*. A one-way ANOVA revealed that there was a tendency for age to differ across groups, $F(2,92)=3.00$, $p=0.55$. However, post hoc analyses of group mean ages using Tukey's method revealed no group differences at the .05 level.

No difference was found across groups in IQ, gross motor skills, fine motor skills, or reading ability (all *F*s *ns*). A MANOVA of perceptual speed, math, and the Preschool Cognitive cluster of the Woodcock Johnson revealed a group difference, Wilks's lambda approximate, $F(6,178)=6.16$, $p<.001$. This was due to math, which was alone in being significant in the univariate analyses, $F(2,91)=5.95$, $p=.004$. Post hoc analyses revealed that this group difference was due to the tutoring group's ($M = 463.80$) performing at a higher level at pretest than either the SIT group ($M = 443.29$) or the perceptual-motor group ($M = 434.06$).

In summary, the samples available for comparison were quite similar at pretest. They differed only in math, for which the children assigned to the tutoring group entered the study at a higher level of performance.

Differential Treatment Effects

MANCOVAs and ANCOVAs were used to test for treatment effects (which would be revealed as group × time interactions) in the three groups (SIT, perceptual-motor, and tutoring), with age as a covariate. For the analyses of academic skills, pretest performance levels were also entered as covariates.

The gross motor and fine motor skills were evaluated from pretest to posttest, but no treatment effects were apparent, $F(4,172) = .43$, *ns*.

An ANCOVA was used to look at treatment effects in reading, but again there was no significant change from pretest to posttest, $F(2,84) = 1.45$, *ns*.

A final MANCOVA was used to evaluate perceptual speed, math, and the Preschool Cognitive Cluster. Although the multivariate result, $F(6,156) = 1.64$, $p = .140$, was not significant at a conventional alpha level, this evaluation was so exploratory that we believed that the result warranted examination of the univariate tests. In addition, we were aware that math was the only variable that revealed any trend toward a significant treatment effect in the Ontario study. Most interesting, in these combined samples it was math that exhibited a significant treatment effect, $F(2,80) = 3.76$, $p = .028$. Post hoc analyses revealed that the tutoring group, which had begun at a higher level of math performance, increased more as a result of treatment than did the SIT or perceptual-motor group. It should be noted that the tutoring group received no extra instruction in math as part of the tutoring program (beyond what they received in the regular classroom).

A final question we asked of the combined data was whether some of the pretest scores were predictive of treatment response. We combined the samples receiving SIT from both studies, for a sample size of 47 (see Table 1). Then we combined the comparison groups, for a sample size of 40 (see Table 2); these were from the Ontario perceptual-motor skills training group and the Alberta tutoring group. In each case, backward elimination multiple regression was performed using all the measures common to both studies. For SIT (see Table 1), the only pretest variables that predicted an improvement on the Woodcock Johnson Reading Cluster were a low pretest Reading Cluster score and a high Preschool Cognitive Cluster score. Even so, R^2 was only .069. Having a low pretest Math Cluster score also predicted improvement on the Math Cluster but accounted for only a small amount of the variance ($R^2 = .095$). Other than a low Perceptual Speed Cluster, the only variable that predicted improvement on this variable was a high pretest Reading Cluster score. Younger children with low Gross Motor scores at pretest tended to improve more than on the Bruininks-Oseretsky Gross Motor Composite. It is interesting that high pretest scores on Gross Motor and high Preschool Cognitive Cluster scores predicted improvement on the Bruininks-Oseretsky Fine Motor Composite.

For the comparison groups (see Table 2), R^2 values were again quite low, ranging from .092 to .248. Improvement in reading was associated with older age at pretest and a low initial score on the Preschool Cognitive Cluster. Improvement in math was related to a low pretest Reading Cluster score. Improvement in Perceptual Speed was associated with a low pretest level in the same variable. Nothing predicted gross motor improvement, but improvements in the Bruininks-Oseretsky Fine Motor Composite were associated with a high initial score in reading and with a younger age at pretest.

The most consistent trend visible in these multiple regression analyses is that improvement in a particular function was associated with a low performance level on that function at pretest. This is consistent with the analyses of each individual study, as reported in Law, Polatajko, Schaffer, Miller, and Macnab (1991) for the Ontario study.

Overall Treatment Effects

Although no difference was found between groups after their treatment, a significant change over time was found in all the groups. Gross motor and fine motor skills changed significantly, $F(2,85) = 7.72$, $p = .001$. In reading, the groups improved significantly, $F(1,85) = 91.08$, $p < .001$. Again, for the grouping of perceptual speed, math, and the Preschool Cognitive Cluster, there was significant change over time, $F(12,207) = 42.25$, $p < .001$. So although we were unable to show the superiority of one type of treatment over another, there was evidence that all groups improved during their 6 months of treatment.

Discussion

Although the Polatajko et al. (1991) study in Ontario and the Wilson et al. (1992) study in Alberta were carried out during overlapping time periods and used many similar dependent measures and selection criteria, the two studies were designed and implemented completely independently. It was only during the course of carrying out the two studies that the investigators learned of each other's activities. When the two studies were completed and the investigators shared their results, the similarity in the outcome, in terms of SIT, was quite apparent to everyone. We then considered whether it would be possible to combine our data to see whether a larger sample

TABLE 1
Backward Elimination Multiple Regression:
Sensory Integration Groups of Both Studies ($n = 47$)

Predictor	Total equation beta	Total equation *p*	Final equation beta	Final equation *p*	*F* (*df*)	R^2
Woodcock-Johnson Reading Cluster					*ns*	.069
Perceptual Speed Cluster	.285	.277				
BOTMP Gross Motor	−.064	.700				
Reading Cluster	−.395	.162	−.429	.094		
Math Cluster	−.022	.926				
Age	−.353	.169				
Preschool Cognitive Cluster	.467	.133	.432	.092		
Woodcock-Johnson Math Cluster					4.722* (1,45)	.095
Perceptual Speed Cluster	−.174	.485				
BOTMP Gross Motor	−.033	.836				
Reading Cluster	.275	.306				
Math Cluster	−.525	.025	−2.173	.035		
Age	−.207	.395				
Preschool Cognitive Cluster	.353	.233				
Woodcock-Johnson Perceptual Speed Cluster					17.43*** (2,44)	.442
Perceptual Speed Cluster	−.821	.000	−.868	.000		
BOTMP Gross Motor	.066	.616				
Reading Cluster	.437	.053	.395	.014		
Math Cluster	−.053	.778				
Age	−.095	.637				
Preschool Cognitive Cluster	.027	.910				
BOTMP—Gross Motor Composite					7.46** (2,44)	.253
Perceptual Speed Cluster	.223	.346				
BOTMP Gross Motor	−.482	.003	−.465	.001		
Reading Cluster	.098	.698				
Math Cluster	−.182	.401				
Age	−.446	.057	−.263	.052		
Preschool Cognitive Cluster	.056	.840				
BOTMP—Fine Motor Composite					5.55***	.201
Perceptual Speed Cluster	.387	.099				
BOTMP Gross Motor	.305	.045	.339	.017		
Reading Cluster	−.225	.366				
Math Cluster	−.373	.083				
Age	−.250	.271				
Preschool Cognitive Cluster	.591	.035	.240	.086		

Note. BOTMP = Bruininks-Oseretsky Test of Motor Proficiency.
*$p < .05$. **$p < .01$. ***$p < .001$.

size would reveal anything more regarding the impact of SIT, and in particular whether any variables were significant predictors of SIT impact. The investigators realized that they had serendipitously used many identical dependent measures, thus facilitating the combining of the research results.

Although one of the primary reasons for combining the two studies was to increase the sample size, thereby compensating, in part, for the known heterogeneity of each sample, this is not an ideal solution. It must be noted that the limitations that are inherent in a single heterogeneous sample still exist for this larger combined sample.

The results of each individual study, and now the results of combining the studies, lead one to the fairly negative

TABLE 2
Backward Elimination Multiple Regression:
Comparison Groups of Both Studies ($n = 40$)

Predictor	Total equation beta	Total equation p	Final equation beta	Final equation p	F (df)	R^2
Woodcock-Johnson Reading Cluster					4.614* (2,37)	.199
Perceptual Speed Cluster	−.054	.832				
BOTMP Gross Motor	.007	.966				
Reading Cluster	−.217	.344				
Math Cluster	.007	.973				
Age	.532	.073	.493	.054		
Preschool Cognitive Cluster	−.580	.063	−.735	.005		
Woodcock-Johnson Math Cluster					7.463** (1,39)	.161
Perceptual Speed Cluster	−.082	.747				
BOTMP Gross Motor	.021	.900				
Reading Cluster	−.286	.171	−.401	.009		
Math Cluster	.371	.111				
Age	−.214	.468				
Preschool Cognitive Cluster	−.122	.695				
Woodcock-Johnson Perceptual Speed Cluster					12.548** (1,38)	.248
Perceptual Speed Cluster	−.539	.029	−.498	.001		
BOTMP Gross Motor	−.081	.614				
Reading Cluster	.236	.275				
Math Cluster	−.033	.868				
Age	−.202	.459				
Preschool Cognitive Cluster	.132	.644				
BOTMP—Gross Motor Composite	No significant betas				N/A	N/A
BOTMP—Fine Motor Composite					*ns*	.092
Perceptual Speed Cluster	.126	.632				
BOTMP Gross Motor	−.144	.409				
Reading Cluster	.452	.041	.451	.023		
Math Cluster	−.130	.587				
Age	−.292	.343	−.372	.057		
Preschool Cognitive Cluster	−.092	.777				

Note. BOTMP = Bruininks-Oseretsky Test of Motor Proficiency.
*$p < .05$. **$p < .01$..

conclusion that SIT is no better than more traditional treatment methods for children with learning or prelearning deficits. These conclusions are limited to the kinds of children used in these two studies. In other words, we have been unable to find any support favoring SIT over other interventions for children with learning or prelearning deficits.

The recent article by Cummins (1991) published in this journal was a fairly scathing analysis of the diagnostic validity of Ayres's work. The current manuscript also has negative implications for Ayres's treatment approach. This leaves one wondering why so many therapists and the families of their clients are still strongly devoted to sensory integration treatment. What is it about SIT that makes people think it is working? We suspect that one factor that might account for this positive interpretation is the intense bond formed between child and therapist. That would certainly explain why the children in the Alberta study who received intensive tutoring improved as much as the SIT group. They, too, developed intense bonds as a result of their one-to-one tutoring program.

Another possible explanation for the popularity of the approach is that therapists and parents *do* perceive that the children improve. Both the Ontario and the Alberta studies found improvements over time as a result of SIT, and these results were maintained when the samples were combined. It is not surprising that clinicians attribute this improvement to SIT (or whatever therapy the child received). What our combined data indicate is that perhaps the attribution of change to SIT is incorrect, not that the perception of change is incorrect. To evaluate whether perceived changes were greater than one could expect from maturation alone will require a no-treatment control group. As mentioned above, the only people who have attempted to implement this control (Polatajko et al., 1991) were unable to continue it because parents were unwilling to enroll their children in a study wherein they might be randomized to a no-treatment control.

Another possible reason for the popularity of an approach whose efficacy has not been shown could be related to the fact that the theoretical framework of sensory integration dysfunction enables the clients and families to "reframe" the motor and behavioral problems ("poorly disciplined, immature, destructive, careless, rigid, or overactive"; Bundy, 1991, p. 319) into terms that are based on a theory that integrates mind and body. When the frame for viewing the client's behavior is changed to a more positive one, the parents and teachers are provided with the basis for developing different kinds of strategies for working with the child. The resulting improvements in behavior, skills, and attitudes are often attributed to the actual physical treatment, not necessarily to any attitudinal change resulting from consultation with the therapist.

It is possible that sensory integration in the 1990s is in the same developmental stage that psychotherapy research was during the 1950s. In 1952, Eysenck published a widely publicized study that questioned the effectiveness of psychotherapy. That article had a powerful effect in stimulating psychotherapy research. Good empirical research began in the 1960s (Hersen, Michelson, & Bellack, 1984). The effectiveness of a wide variety of types of psychotherapeutic approaches is now well established, but only after 40 years of intensive research with increasingly rigorous methodologies. A significant portion of this research literature has investigated process variables, patient characteristics, and therapist characteristics. Although SIT researchers have begun to consider many of these variables, they have a long road to travel before its effectiveness is established or refuted.

ABOUT THE AUTHORS

Bonnie J. Kaplan *is professor of pediatrics and psychology at the University of Calgary. Her research interests are physiological correlates (including sensorimotor deficits) of childhood learning and behavior disorders.* ***Helene J. Polatajko*** *is associate professor of occupational therapy and educational psychology at The University of Western Ontario. She has published numerous articles on the motor and sensory aspects of learning disabilities in children and adults. Her research has focused on assessment and intervention issues.* ***Brenda N. Wilson*** *is an occupational therapist and research coordinator/consultant at the Alberta Children's Hospital Research Centre. She received her graduate degree from Boston University. Her research interests are identification, assessment, and treatment of the sensorimotor problems of children with learning disabilities.* ***Peter D. Faris*** *is a graduate of the University of Calgary. His research interests are experimental design and methodology. Address: Bonnie J. Kaplan, Department of Pediatrics, Alberta Children's Hospital Research Centre, 1820 Richmond Road SW, Calgary, Alberta, Canada T2T 5C7.*

AUTHORS' NOTES

Preparation of this article was supported by the Alberta Children's Hospital Foundation and an Alberta Mental Health Research Fund Visiting Scientist Award to Dr. Polatajko.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Ayres, A.J. (1972). Improving pediatric scores through sensory integration. *Journal of Learning Disabilities, 5,* 24–28.

Bruininks, R.H. (1978). *Bruininks-Oseretsky test of motor proficiency, examiner's manual.* Circle Pines, MN: American Guidance Service.

Bundy, A.C. (1991). Consultation and sensory integration theory. In A.G. Fisher, E.A. Murray, & A.C. Bundy (Eds.), *Sensory integration theory and practice* (pp. 318–332). Philadelphia: F.A. Davis.

Cummins, R.A. (1991). Sensory integration and learning disabilities: Ayres' factor analyses reappraised. *Journal of Learning Disabilities, 24,* 160–168.

Eysenck, H.J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16,* 319–324.

Hersen, M., Michelson, L., & Bellack, A.S. (1984). *Issues in psychotherapy research.* New York: Plenum Press.

Law, M., Polatajko, H., Schaffer, R., Miller, J., & Macnab, J. (1991). The impact of heterogeneity in a clinical trial: Motor outcomes after sensory integration therapy. *Occupational Therapy Journal of Research, 11,* 117–189.

Ottenbacher, K. (1982). Sensory integration therapy: Affect or effect. *The American Journal of Occupational Therapy, 36,* 571–577.

Polatajko, H., Kaplan, B.J., & Wilson, B.N. (in press). Sensory integration treatment for children with learning disabilities: Its status 20 years later. *Occupational Therapy Journal of Research.*

Polatajko, H., Law, M., Miller, J., Schaffer, R., & Macnab, J. (1991). The effect of a sensory integration program on academic achievement, motor performance and self-esteem in children identified as learning disabled: Results of a clinical trial. *The Occupational Therapy Journal of Research, 11,* 155–176.

Schaffer, R., Law, M., Polatajko, H., & Miller, J. (1989). A study of children with learning disabilities *or* Let's not throw the baby out with the bathwater. *Physical & Occupational Therapy in Pediatrics, 9,* 101–117.

Slosson, R.L. (1963). *The Slosson intelligence test for children and adults.* New York: Slosson Educational Publications.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children–Revised.* San Antonio, TX: Psychological Corp.

Wilson, B.N., Kaplan, B.J., Fellowes, S., Gruchy, C., & Faris, P. (1992). The efficacy of sensory integration treatment compared to tutoring. *Physical and Occupational Therapy in Pediatrics, 12,* 1–36.

Woodcock, R.W., & Johnson, M.B. (1977). *Woodcock-Johnson psycho-educational battery.* Hingham, MA: Teaching Resources Corp.