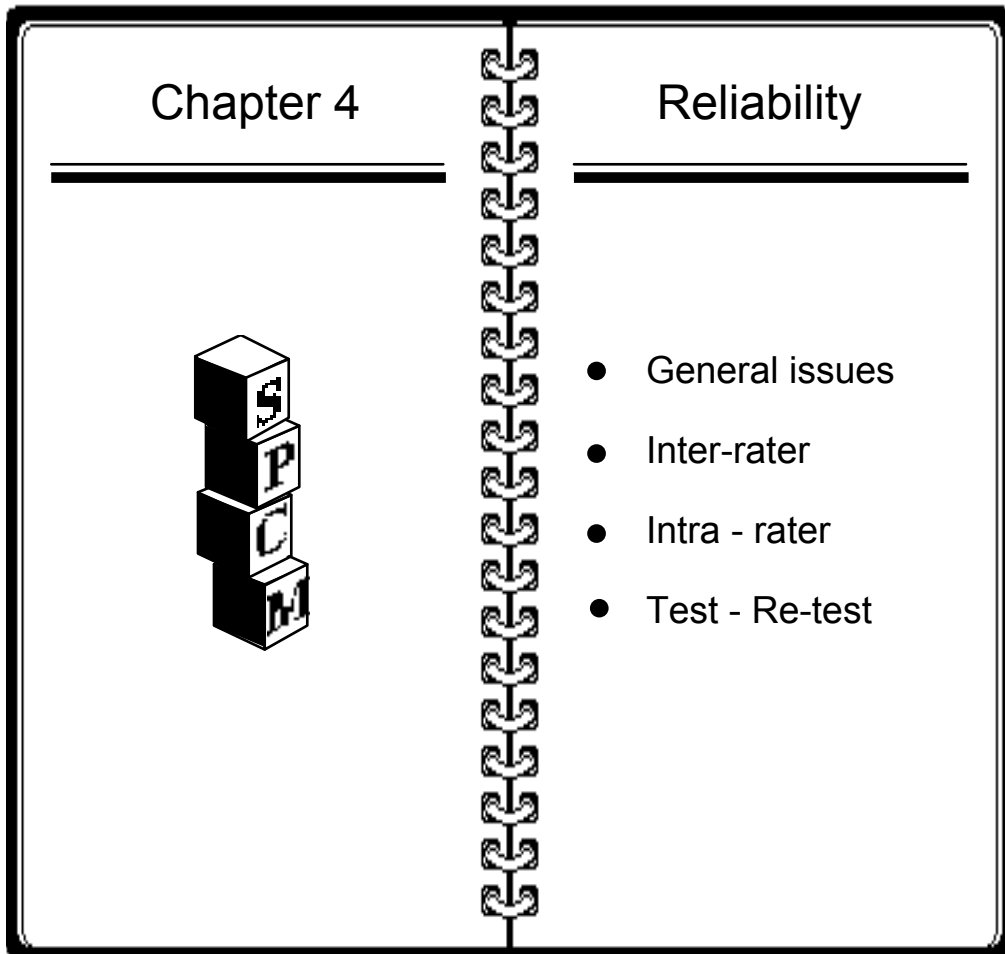


Chapter 4A: Reliability – Trial 1



Following the development process outlined in Chapter 3, two clinical trials were conducted, both with the primary purpose of determining reliability of the SPCM. In both trials, inter-rater and test-retest reliability were examined. In the first trial, Trial 1, subjects were administered the SPCM both while sitting 'with' and 'without' their own seating systems while in Trial 2, subjects were examined only in the 'with' condition.

Trial protocols were approved by the ethical screening committee of the University of British Columbia.

Trial 1

Results of the first clinical trial, conducted in September-November, 1990, were presented in two reports (Fife, Roxborough, Armstrong, Harris, Gregson, Field, 1991; Fife, 1992).

Subjects: Trial 1

The subject sample of 45 children included all clients of the Positioning Assessment Unit at Sunny Hill Hospital who were from the Greater Vancouver Regional District, were less than 19 years of age and for whom informed consent had been obtained. All subjects were non-ambulatory and users of seating systems prescribed prior to initiation of the trial. The sessions with the first four subjects were scheduled as practice trials and their scores were not included in the data analysis. An additional subject did not participate due to her inability to attend during scheduled session times. Thus, 40 subjects participated in Trial 1.

Raters: Trial 1

The two raters were an occupational therapist and physical therapist (D.F. and J.G.), each with at least five years' experience in pediatrics including two or more years in the adaptive seating field. In addition to administering the SPCM in practice trials with the first four subjects, the raters had informally administered some of the items in their regular practice.

Procedures: Trial 1

Data were collected over a six-week period. Subjects attended two assessment sessions which were approximately three weeks apart. We assumed first that a three-

week interval would be adequate for the raters to forget scores from the first assessment and, second, that the children would not undergo significant clinical change between sessions.

Both the LSS and the SPCM were administered to subjects at each of the two assessment sessions. Each rater independently administered the LSS while the other rater was outside of the testing room. Then the SPCM was administered by each rater to subjects in two seating conditions: 'with' and 'without' their seating systems. The rationale for administering the SPCM under both seating conditions was that we propose to use the SPCM to compare alignment and function outcomes under both conditions.

If children were not independent sitters, i.e., if they scored less than 5 on the LSS, they were provided with Tumbleform Feeder Seats¹ for the 'without' condition. This procedure was done to avoid using 'hands-on' support which is difficult to standardize. The Tumble-form seats with a lap belt were placed on a modified wheeled base² which could be reproducibly oriented in space. The raters jointly selected an orientation of the base judged to be optimal for the safety, comfort and function of the individual subject. If subjects scored 5 or more on the LSS, they sat on a bench with feet supported while the SPCM was administered in the 'without' seating system condition.

The order of testing alignment vs function for each subject was alternated in each condition. The order of testing conditions, i.e., 'with' vs 'without' seating system, was alternated in successive subjects. For each subject, this sequence was maintained in the second session. Both raters assessed and independently scored the alignment items simultaneously. One rater administered the function items while the other rater observed and each scored the items independently. The same rater administered the function items to a given subject at both sessions and the administering rater was alternated for successive subjects. Subjects attended in pairs for sessions of approximately two hours duration. This allowed for rest periods and repositioning of subjects between administration of the LSS and the SPCM under the two test conditions.

¹JA Preston Corp., 60 Page Road, Clifton, NJ 07012

²M.O.S.S. II Modular Orthotic Seating System, Otto Bock Orthopedic Industry, 4120 Highway 55, Minneapolis, MN 55422

Data Analysis: Trial 1

Items

Inter-rater and test-retest reliability for the LSS and SPCM items were assessed by examination of agreement tables. The percent of agreement ($\#$ of agreements/total observations \times 100) and the Kappa statistic were used as indices of estimated reliability. The Kappa statistic reflects the percent agreement over and above chance.^{43,44} Kappa values will be low, even if percent agreement is high, if the observed scores for an item fall mainly in one category. In this case, there is a greater probability that chance alone is responsible for agreement. Haley and colleagues, in assessing reliability of the Movement Assessment of Infants,⁴⁰ interpreted Kappa values as follows: Poor, if less than .40; Fair to Good, if .40 to .74; and Excellent, if .75 or higher.⁴⁵ We chose a Kappa value of .40 as our minimum level of acceptable reliability. The percent agreement statistic is included in the table because of its frequent use in the literature. However, decision-making was based on the Kappa results because this statistic corrects for chance.

Component Scores

Six SPCM component scores were calculated: summed item scores for each domain, alignment and function; and component alignment scores for the head, trunk, pelvis and lower limbs, each calculated as the sum of all alignment item scores for the respective body segment. Each SPCM component score was tested for inter-rater and test-retest reliability using intra-class correlation coefficients (Shrout and Fleiss, 1979).

Results: Trial 1

Subjects are described in Tables I and 2. The mean age was 9.1 years, with a range of 1.7 to 18.5 years. Almost half of the subjects (19) were in the diagnostic group, cerebral palsy.

Table 1. Characteristics of Subjects in Trial 1 (N=40)

Variable	N
Sex	
Male	22
Female	18
Age Category (yr)	
<3	6
3-5	8
6-9	5
Medical Diagnosis	
Cerebral Palsy	
Spastic quadriplegia	15
Diplegia	2
Triplegia	1
Hemiplegia	1
Brain injury	7
Meningomyelocele	2
Muscle disease	3
Developmental delay	2
Other	7
Ability to understand verbal instructions (opinion of rater)	
Yes	19
No	21

Level of Sitting Scale

Descriptors of subject sitting independence, as defined by the Level of Sitting Scale (described in Chapter 6) were spread over the seven scale levels from 0, 'unplaceable', to 6, 'shifts trunk laterally, re-erects'. (Note: The seven scale levels for Trial 1 were numbered 0 to 6. In Trial 2 the same descriptors were used but the score range was re-numbered from I to 7.) As seen in Table 2, the most common descriptor for Trial I subjects was Level 2, 'requires support from shoulders or trunk downward'.

Table 2. Distribution of LSS scores in Trial 1 subjects

Level	Test 1 a		Test2 b	
	Rater 1	Rater 2	Rater 1	Rater 2
0	2	4	3	4
1	2	2	4	2
2	19	14	11	10
3	3	3	2	6
4	6	9	8	6
5	3	1	3	2
6	3	5	4	5

^a n = 38 subjects

^b n = 35 subjects

The mean Kappa value for inter-rater reliability over the two test sessions was .60, with mean agreement of 69%. The Kappa mean for test-retest reliability across the two raters was .55, with mean agreement of 64%. These results are presented in Table 3.

Table 3. Reliability of Level of Sifting Scale in Trial 1

<i>Inter-Rater</i>	No. of Subjects	Kappa	Percent
Test 1	38	0.58	68.4
Test 2	35	0.62	68.6
<i>Test-Retest</i>			
Rater 1	34	0.54	64.7
Rater 2	33	0.55	63.6

Items

With respect to inter-rater reliability across the two seated conditions and the two test sessions three weeks apart, the overall mean of the item Kappa coefficients was .45 for the alignment section and .85 for the function section. Kappa coefficients and percent agreement for individual items in each of the four data sets are presented in Table 4. Within the four data sets, the number of alignment items with Kappa values under the acceptable level of .40 ranged from 6 to 12 out of a possible 22 items. In the function section, one of 12 items (# 1) had a Kappa value less than .40. There were five alignment and 11 function items with .40 or better Kappa values across all four data sets.

Table 4. Inter-rater reliability of SPCM items in Trial 1: Kappa values and percent agreement

Item Description	With Seat Test 1		With Seat Test 2		Without Seat Test 1		Without Seat Test 2	
	Kappa	%	Kappa	%	Kappa	%	Kappa	%
<i>Alignment Items</i>								
A1 Pelvic obliquity	0.52	73.7	0.44	71.4	0.36	63.2	0.31	57.1
A2 Trunk lat incline	0.51	83.8	0.84	94.3	0.52	70.3	0.39	67.6
A3 Shoulder height	0.35	65.8	0.30	62.9	0.42	70.3	0.46	73.5
A4 Head lat tilt	0.52	77.8	0.60	82.9	0.43	67.6	0.54	76.5
A5 Hip R rotation	0.24	57.9	0.37	65.7	0.60	77.8	0.58	77.1
A6 Hip L rotation	0.66	81.6	0.54	74.3	0.75	86.1	0.68	82.9
A7 Pelvic tilt	0.55	73.7	0.32	62.9	0.31	50.0	0.26	48.5
A8 Lumbar curve	0.08	48.6	0.54	77.1	-0.02	32.4	0.06	37.1
A9 Thoracic curve	0.28	57.9	0.41	71.4	0.30	56.8	0.26	54.3
A10 Trunk AP incline	0.49	73.0	0.50	74.3	0.59	73.7	0.52	71.4
All Head AP tilt	0.24	66.7	0.46	73.5	0.39	60.5	0.31	57.6
A12 Hip R fl/ext	0.42	80.6	0.77	94.3	0.37	68.4	0.26	62.9
A13 Hip L fl/ext	0.31	75.0	0.77	94.3	0.53	76.3	0.33	65.7
A14 Knee R fl/ext	0.37	78.9	0.62	91.4	0.76	89.5	0.75	91.4
A15 Knee L fl/ext	0.42	81.6	0.52	91.4	0.47	76.3	0.62	85.7
A16 Ankle R df/pf	0.31	89.5	1.00	100.0	0.73	86.8	0.70	88.6
A17 Ankle L df/pf	0.25	86.5	0.87	97.1	0.77	89.5	0.61	85.7
A18 Pelvic rotation	0.43	67.6	0.30	61.8	0.20	56.8	0.29	57.1
A19 Trunk rotation	0.49	76.3	0.12	60.0	0.07	52.6	-0.05	42.9
A20 Head rotation	0.59	86.5	0.61	88.6	0.53	77.8	0.44	77.1
A21 Hip R add/abd	0.47	69.4	0.53	73.5	0.63	77.1	0.36	58.8
A22 Hip L add/abd	0.32	59.5	0.21	54.3	0.67	81.1	0.27	55.9
<i>Mean</i>	0.40	73.3	0.53	78.1	0.47	70.0	0.41	67.1
Mean across all data sets	Kappa: 0.45		Percent Agreement: 72.1					
<i>Function Items</i>								
F1 Head up AP	0.16	89.5	0.16	74.3	0.46	84.2	0.21	68.6
F2 Head up midline	0.90	86.8	0.90	94.3	0.54	70.3	0.82	88.2
F3 Trunk move AP	0.85	91.9	0.85	91.2	0.81	88.9	0.91	94.3
F4 Trunk move rot	0.95	87.9	0.95	97.1	0.79	88.2	0.87	91.4
F5 Arm lift	0.92	92.1	0.92	94.3	0.57	67.6	0.89	91.4
F6 Grasp block	0.95	97.4	0.95	97.1	0.92	94.6	1.00	100.0
F7 Grasp raisin	0.96	97.3	0.96	97.1	0.88	91.9	1.00	100.0
F8 Manipulate jar	1.00	94.7	1.00	100.0	0.84	89.2	0.92	94.3
F9 Manipulate pen	0.89	100.0	0.89	94.1	0.89	94.3	1.00	100.0
F10 Manipulate dice	0.95	94.7	0.95	97.1	0.90	94.4	1.00	100.0
F11 WC mobil. speed	1.00	94.4	1.00	100.0	1.00	100.0	1.00	100.0
F12 WC mobil. accur.	1.00	100.0	1.00	100.0	0.91	96.9	0.79	92.6
<i>Mean</i>	0.88	93.9	0.88	94.7	0.79	88.4	0.87	93.4
Mean across all data sets	Kappa: 0.85		Percent Agreement: 92.6					

With respect to test-retest reliability across the two seated conditions and two raters, the overall mean of the item Kappa coefficients was .35 for the alignment section and .29 for the function section. Kappa coefficients and percent agreement for individual items in each of the four data sets are presented in Table 5.

Table 5. Test-retest reliability of SPCM items in Trial 1: Kappa values and percent agreement

Item	Description	<i>With Seat</i>		<i>With Seat</i>		<i>Without Seat</i>		<i>Without Seat</i>	
		<i>Rater 1</i>	<i>Rater 2</i>	<i>Rater 1</i>	<i>Rater 2</i>	<i>Rater 1</i>	<i>Rater 2</i>	<i>Rater 1</i>	<i>Rater 2</i>
		Kappa	%	Kappa	%	Kappa	%	Kappa	%
<i>Alignment Items</i>									
A1	Pelvic obliquity	0.16	57.1	0.43	68.6	0.08	45.5	0.11	42.4
A2	Trunk lat incline	0.77	91.2	0.63	88.6	0.18	47.1	0.35	63.6
A3	Shoulder height	0.44	71.4	0.44	74.3	0.44	67.6	0.30	66.7
A4	Head lat tilt	0.41	72.7	0.51	79.4	0.35	63.6	0.41	69.7
A5	Hip R rotation	0.35	62.9	0.02	48.6	0.44	68.8	0.32	64.7
A6	Hip L rotation	0.29	60.0	0.07	48.6	0.27	59.4	0.26	58.8
A7	Pelvic tilt	0.22	57.1	0.36	62.9	0.33	53.3	0.16	40.6
A8	Lumbar curve	0.00	58.8	0.32	60.0	0.15	44.1	0.14	41.2
A9	Thoracic curve	0.24	65.7	0.23	54.3	0.20	48.5	0.46	64.7
A10	Trunk AP incline	0.39	67.6	0.45	71.4	0.48	67.6	0.77	85.3
All	Head AP tilt	0.44	72.7	0.55	79.4	0.50	67.6	0.50	72.7
A12	Hip R fl/ext	0.12	70.6	0.46	85.7	0.59	79.4	0.61	79.4
A13	Hip L fl/ext	0.09	67.6	0.40	82.9	0.70	85.3	0.61	79.4
A14	Knee R fl/ext	0.53	85.7	0.31	82.9	0.62	85.3	0.60	82.4
A15	Knee L fl/ext	0.68	91.4	0.37	85.7	0.46	79.4	0.62	82.4
A16	Ankle R df/pf	0.21	85.7	0.48	94.3	0.50	76.5	0.66	85.3
A17	Ankle L df/pf	0.10	79.4	0.37	91.4	0.66	85.3	0.70	88.2
A18	Pelvic rotation	0.28	60.6	0.13	54.5	0.16	47.1	0.17	53.1
A19	Trunk rotation	0.04	51.4	0.16	65.7	-0.03	44.1	0.17	61.8
A20	Head rotation	0.24	73.5	0.19	74.3	0.18	61.8	0.45	76.5
A21	Hip R add/abd	0.25	55.9	0.34	60.0	0.39	57.6	0.23	52.9
A22	Hip L add/abd	0.56	74.3	0.25	58.8	0.39	60.6	0.34	61.8
<i>Mean</i>		0.31	69.7	0.34	71.5	0.36	63.4	0.41	67.0
<i>Mean across all data sets</i>		<i>Kappa: 0.35</i>				<i>Percent Agreement: 67.9</i>			
<i>Function Items</i>									
F1	Head up AP	-0.15	3.2	-0.03	2.9	-0.21	7.4	-0.03	0.0
F2	Head up midline	-0.08	35.5	-0.11	34.4	0.03	40.6	-0.13	32.3
F3	Trunk move AP	0.41	60.0	0.32	51.5	0.33	50.0	0.25	40.6
F4	Trunk move rot	0.42	65.7	0.20	46.7	0.28	53.1	0.25	45.2
F5	Arm lift	0.24	45.7	0.37	54.3	0.15	35.5	0.45	59.4
F6	Grasp block	0.33	50.0	0.35	50.0	0.26	42.4	0.25	41.2
F7	Grasp raisin	0.29	42.9	0.31	44.1	0.17	30.3	0.19	29.4
F8	Manipulate jar	0.36	57.1	0.37	57.1	0.23	46.9	0.43	60.6
F9	Manipulate pen	0.38	57.1	0.31	51.5	0.23	46.9	0.27	45.5
F10	Manipulate dice	0.34	60.0	0.32	60.0	0.38	60.6	0.44	63.6
F11	WC mobil. speed	0.63	82.4	0.59	81.3	0.77	92.6	0.77	92.3
F12	WC mobil. accur.	0.49	78.1	0.49	77.4	0.72	92.3	0.32	77.8
<i>Mean</i>		0.31	53.1	0.29	50.9	0.28	49.9	0.29	49.0
<i>Mean across all data sets</i>		<i>Kappa: 0.29</i>				<i>Percent Agreement: 50.7</i>			

Within the four data sets, the number of alignment items with Kappa values under the acceptable level of .40 ranged from 11 to 15 out of a possible 22 items. The number of function items with Kappa values less than .40 ranged from 8 to 10 out of a possible 12 items. Only one alignment and one function item had a test-retest Kappa value of .40 or better across all four data sets.

Component Scores

Mean SPCM component scores with and without the seating system are shown in Table 6.

Table 6. Descriptive Statistics of SPCM Component Scores in Trial 1

SPCM Component	N ^c	Possible Score ^b		With Seat		Without Seat	
		Min	Max	Mean	sd	Mean	sd
<i>Alignment</i>	22	0	66	8.6	5.5	14.0	6.6
Pelvis	3	0	9	1.7	1.6	3.1	1.6
Trunk	6	0	18	2.5	2.1	4.6	2.4
Head	3	0	9	1.0	1.4	1.7	1.9
Lower Limbs	10	0	30	3.3	2.3	4.7	2.9
Function	12	0	36	16.0	12.1	14.2	11.1

a Mean of mean scores and standard deviations in 50 subjects across four data sets: Rater 1:Test 1, Rater 2:Test 1, Rater 1:Test 2, Rater 2:Test 2

b Note that item scores ranged from 0 to 3. In the alignment domain, the higher the score, the greater the deviation from normal alignment with 0 representing normal alignment, and 1, 2, and 3 representing respectively, mild, moderate and severe deviations from normal alignment. In re-coding from the data sheets, right and left designations of deviations were ignored, i.e., it was assumed that the raters correctly distinguished right from left-sided deviations and had recorded this accurately. In the function domain, the higher the score, the better the performance of the item.

c Number of alignment items comprising SPCM component

Intraclass correlation coefficients (ICCs) of SPCM component scores ranged from .49 to 1.00 (Table 7). The mean inter-rater reliability ICC across the four data sets for total scores of alignment and function domains was .87 and .98, respectively and the corresponding means for test-retest reliability were .83 and .95 respectively.

Table 7. Reliability^a of SPCM component scores in Trial 1

Component	N ^b	C ^c 1 T ^d 1	Inter-rater			Test-Retest			
			C1 T2	C2 T1	C2 T2	C1 R ^e 1	C1 R2	C2 R1	C2 R2
<i>Alignment</i>	22	.82	.88	.90	.87	.78	.83	.85	.84
Pelvis	3	.86	.75	.78	.55	.69	.67	.56	.52
Trunk	6	.74	.78	.67	.71	.55	.67	.57	.62
Head	3	.84	.87	.85	.82	.80	.57	.80	.81
Lower Limbs	10	.49	.69	.77	.75	.53	.64	.80	.71
Function	12	1.00	.99	.94	.99	.98	.96	.90	.97-

^a Intraclass correlation coefficients, formula 3, 1, as described by Shrout and Fleiss (1979)

^b N = number of items in SPCM Component

^c C = Condition (C 1: 'with seating system'; C 2: 'without seating system')

^d T = Test (T 1: Test 1; T 2: Test 2)

^e R = Rater (R 1: Rater 1; R 2: Rater 2)

Discussion: Trial I

While reliability of the SPCM alignment and function domain scores was satisfactory both with and without the subject being in a seating system, it was considered worthwhile to attempt to improve reliability for the summed alignment segment scores and for individual items.

It was decided to retain all SPCM items but descriptions would be modified to improve clarity. For the alignment domain, the category widths, i.e., the range of postural deviation within normal, mild, moderate and severe categories for each item, would be adjusted to better reflect the clinical judgement of two seating clinicians, one of whom took part in Trial 1. In most cases, the category widths were decreased. The numerical scores for categories would be changed from 0 to 3, as in Trial 1, to 1 to 4. The use of two visual aids would be introduced to enhance visual observation of postural alignment: a flexible curve and a protractor or goniometer (see Chapter 6 for details). Finally, it was suggested that accuracy of the raters improved with practice. SPCM administration to 10 clients was recommended by the occupational therapist and physiotherapist participating as raters in this trial.