

PCIT Reliability
Grade 1/Year 2
Fast Track Project Technical Report
Cynthia Rains
April 2, 2003

Table of Contents

- I. Scale Description
- II. Report Sample
- III. Scaling
- IV. Differences Between Groups
- V. Recommendations for Use
- VI. Means and SDs
- VII. Correlations

Appendix

SAS scoring program

Citations

References

Conduct Problems Prevention Research Group (CPPRG). (1993). *Parent-child interaction task: observational data collection manuals*. Unpublished manuscript, Fast Track Project.

Crnic, K. and Greenberg, M. (1990). Minor parenting stresses with young children. *Child Development*, 61, 1628-1637.

Forehand, R. & McMahon, R. (1981). *Helping the noncompliant child: A clinician's guide to parent training*. New York: Guilford Press.

Report

Rains, C. (2003). *PCIT Reliability* (Fast Track Project Technical Report). Available: <http://www.fasttrackproject.org>

Data Sources

Raw: 02Erel

Scored: PCR2

I. Scale Description

The Parent-Child Interaction Task (PCIT) was administered to the child and the parent as part of the summer interview in the child's home. Two measures are included in the PCIT: the Behavioral Coding System (BCS; Forehand & McMahon, 1981) and the Interaction Ratings Scale (Crnic and Greenberg, 1990). The PCIT is comprised of four separate segments, which occur in the following sequence: Child's Game (free play, 5 minutes), Parent's Game (parent-directed play, 5 minutes), Lego Task (5 minutes), and Clean-Up (3 minutes). The BCS is used by the child interviewer to record both parent behaviors and child behaviors during each task; the amount of time for each task and the total time are also recorded and used in the scoring of the subscales. The child interviewer then completes the IRS, which is a set of 16 global ratings of the mother and child with respect to gratification, sensitivity, and involvement, after each of the 4 tasks that the parent and child complete together.

To check the reliability of the scoring by the child interviewer, the site's lead interviewer or assistant lead interviewer accompanied the assigned pair of interviewers to complete a version of the BCS and the IRS. These scores were then compared to the child interviewer's scores. For

the BCS variables, the subscales (commands, positive attention, negative attention, compliance, noncompliance, and disruptive behavior) were compared for percent agreement. For the IRS variables, three things were checked: the number of items with a reliability of zero, the sum of all of the reliability ratings, and the number of items with exact agreement. These reliability checks were performed on 15% of the home visits. Reliability checks were supposed to be distributed evenly across interviewers but not across samples (intervention, control, normative).

II. Report Sample

These exploratory analyses were conducted with the first cohort on the control sample (n=155) and on the normative sample (n=387, 463 with overlap) during the second year of the study. Overall, 81 records were included in the reliability check. 22 of these were from the control sample (7 from Durham, 6 from Nashville, 5 from Pennsylvania, and 4 from Washington) and 68 were from the normative sample (19 from Durham, 16 from Nashville, 18 from Pennsylvania, and 15 from Washington). Of these, only a small number had a reliability check for all of the scores: 8 were from the control sample (1 from Durham, 1 from Nashville, 2 from Pennsylvania, and 4 from Washington) and 18 were from the normative sample (1 from Durham, 1 from Nashville, 2 from Pennsylvania, and 14 from Washington). A number of records were missing anywhere from 1 to 6 reliability checks.

III. Scaling

Separate methods were used for the BCS and IRS to calculate the reliability checks.

For the BCS subscales, in order to obtain a percentage interobserver agreement, the score sheet of the child interviewer is compared with that of the lead observer. To calculate percentage interobserver agreement for 5 of the BCS scales subscales (commands, positive attention, negative attention, compliance, noncompliance), the following steps are taken:

- 1) Calculate the number of agreements and disagreements between the lead observer and the child interviewer for each 30-second interval. Agreements are defined at the level of total number of behaviors scored in the interval, rather than the specific location of the behaviors within the interval.
- 2) Then, the total number of agreements for a behavior (positive attention, compliance, etc.) across the intervals is summed.
- 3) Next, the total number of disagreements for a behavior across the intervals is summed.
- 4) Last, the percentage is calculated by dividing the number of agreements by the sum of the number of agreements and the number of disagreements.
- 5) The same procedure is used for each of these five scales (behaviors) for the BCS.

A slightly different procedure is used for the BCS scale, Child Disruptive Behavior. There are 36 possible occurrences of this type of behavior. In order to calculate the reliability, only intervals where both the child interviewer and the lead observer coded this behavior are used. Then, the number of agreements and disagreements are summed, followed by using the same percentage formula as with the other BCS scales.

Finally, a Total Percentage Interobserver Agreement is calculated by summing the number of agreements for all six behaviors, then summing the number of disagreements for all six behaviors, and using the same percentage formula, as mentioned above.

For any percentage agreement for the BCS scales, reliability is considered adequate if it is 70% or higher, per the expectations of the researchers.

To check the reliability for the IRS, the lead observer again compares his/her score sheet with the score sheet of the child interviewer. Each pair of scores is to receive a score as follows:

- 0 = scores match exactly or are ± 1
- 1 = scores that differ by ± 2
- 2 = scores that differ by ± 3

The researchers set values for reliability levels for the IRS. 12 out of 16 points or 75% of the ratings should have a reliability of zero. This is to ensure a reasonable degree of accuracy. The sum of all the IRS reliability ratings should not exceed 4 to ensure that there is no excessive variability in the ratings.

IV. Differences between Group

A series of t-tests between the high-risk control sample and the normative sample (including the overlap) indicated no significant differences for any of the BCS scales and IRS variables.

PCIT Reliability

Variable	Normative Sample		Control Sample		DF	t Value	Pr > t
	Mean	Std Dev	Mean	Std Dev			
BCS: COMMANDS % AGREEMENT (O2E_R1)	81.12	11.93	80.36	12.90	79	-0.25	0.8049
BCS: POSITIVE ATTENTION % AGREEMENT (O2E_R2)	75.83	20.92	64.37	30.65	74	-1.83	0.0716
BCS: NEGATIVE ATTENTION % AGREEMENT (O2E_R3)	63.77	46.04	72.25	34.15	32	0.56	0.5807
BCS: COMPLIANCE % AGREEMENT (O2E_R4)	73.14	16.75	72.18	14.25	79	-0.24	0.8134
BCS: NONCOMPLIANCE % AGREEMENT (O2E_R5)	70.08	40.28	73.20	35.73	50	0.26	0.7953
BCS: DISRUPTIVE BEHAVIOR % AGREEMENT (O2E_R6)	75.33	35.01	83.23	23.85	38	0.73	0.4680
BCS: TOTAL % AGREEMENT (O2E_R7)	78.09	11.39	76.77	10.55	79	-0.47	0.6396
IRS: # OF ITEMS WITH RELIABILITY OF ZERO (O2E_R8)	15.39	0.98	15.52	0.75	78	0.57	0.5718
IRS: SUM OF RELIABILITY RATINGS (O2E_R9)	0.61	0.98	0.52	0.81	78	-0.36	0.7193
IRS: # OF ITEMS WITH EXACT AGREEMENT (O2E_R10)	8.75	2.75	8.86	2.59	78	0.16	0.8718

V. Recommendations for Use

Analysts should note that distribution properties for these scores might be nonnormal due to the small sample sizes involved and the nature of the variable construction. With that in mind, Commands % Agreement, Disruptive Behavior % Agreement, Total % Agreement, and Number of Items with Reliability of Zero were all negatively skewed for both the normative and the control samples.

Another approach to reliability would have been to calculate the inter-class correlations for these two measures. However, raw data for the lead observer were not available to perform these calculations.

VI. Means and SDs

PCIT Reliability Normative Sample Year 2

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
O2E_r1	BCS: Commands % Agreement	68	80.0735294	12.6435827	33.0000000	100.0000000
O2E_r2	BCS: Positive Attention % Agreement	64	74.2656250	21.2238576	0	100.0000000
O2E_r3	BCS: Negative Attention % Agreement	27	62.4444444	44.8487487	0	100.0000000
O2E_r4	BCS: Compliance % Agreement	68	72.0000000	16.7955218	25.0000000	100.0000000
O2E_r5	BCS: Noncompliance % Agreement	43	69.0232558	39.6084938	0	100.0000000
O2E_r6	BCS: Disruptive Behavior % Agreement	33	77.5454545	32.9697692	0	100.0000000
O2E_r7	BCS: Total % Agreement	68	77.0735294	11.6264189	31.0000000	100.0000000
O2E_r8	IRS: # of Items with Reliability of 0	68	15.3823529	0.9624194	12.0000000	16.0000000
O2E_r9	IRS: Sum of All Reliability Ratings	68	0.6323529	0.9758947	0	4.0000000
O2E_r10	IRS: # of Items with Exact Agreement	68	8.5294118	2.6736447	2.0000000	16.0000000

PCIT Reliability Control Sample Year 2

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
O2E_r1	BCS: Commands % Agreement	22	80.3636364	12.8971959	34.0000000	100.0000000
O2E_r2	BCS: Positive Attention % Agreement	19	64.3684211	30.6525521	0	100.0000000
O2E_r3	BCS: Negative Attention % Agreement	12	72.2500000	34.1470749	0	100.0000000
O2E_r4	BCS: Compliance % Agreement	22	72.1818182	14.2482626	45.0000000	100.0000000
O2E_r5	BCS: Noncompliance % Agreement	15	73.2000000	35.7335377	0	100.0000000
O2E_r6	BCS: Disruptive Behavior % Agreement	13	83.2307692	23.8542583	33.0000000	100.0000000
O2E_r7	BCS: Total % Agreement	22	76.7727273	10.5466394	45.0000000	95.0000000
O2E_r8	IRS: # of Items with Reliability of 0	21	15.5238095	0.7496031	14.0000000	16.0000000
O2E_r9	IRS: Sum of All Reliability Ratings	21	0.5238095	0.8135753	0	2.0000000
O2E_r10	IRS: # of Items with Exact Agreement	21	8.8571429	2.5939490	5.0000000	14.0000000

VII. Correlations

PCIT Reliability Report Sample Year 2

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations					
	O2E_r1	O2E_r2	O2E_r3	O2E_r4	O2E_r5
O2E_r1 BCS: Commands % Agreement	1.00000 81	0.16649 0.1506 76	0.34467 0.0459 34	0.42521 <.0001 81	0.30354 0.0287 52
O2E_r2 BCS: Positive Attention % Agreement	0.16649 0.1506 76	1.00000 76	0.11996 0.4992 34	0.08700 0.4549 76	0.25760 0.0739 49
O2E_r3 BCS: Negative Attention % Agreement	0.34467 0.0459 34	0.11996 0.4992 34	1.00000 34	0.24918 0.1553 34	0.55389 0.0022 28
O2E_r4 BCS: Compliance % Agreement	0.42521 <.0001 81	0.08700 0.4549 76	0.24918 0.1553 34	1.00000 81	0.39432 0.0038 52
O2E_r5 BCS: Noncompliance % Agreement	0.30354 0.0287 52	0.25760 0.0739 49	0.55389 0.0022 28	0.39432 0.0038 52	1.00000 52
O2E_r6 BCS: Disruptive Behavior % Agreement	-0.05025 0.7581 40	-0.00922 0.9556 39	-0.04541 0.8293 25	0.12797 0.4313 40	0.17673 0.3025 36
O2E_r7 BCS: Total % Agreement	0.90666 <.0001 81	0.33374 0.0032 76	0.41785 0.0139 34	0.66466 <.0001 81	0.43661 0.0012 52
O2E_r8 IRS: # of Items with Reliability of 0	0.15854 0.1601 80	-0.00809 0.9447 76	0.36941 0.0315 34	0.17996 0.1102 80	-0.00564 0.9687 51
O2E_r9 IRS: Sum of All Reliability Ratings	-0.16293 0.1487 80	-0.01001 0.9316 76	-0.38351 0.0252 34	-0.19675 0.0803 80	-0.00498 0.9724 51
O2E_r10 IRS: # of Items with Exact Agreement	-0.02093 0.8538 80	-0.07427 0.5237 76	0.05500 0.7574 34	-0.04714 0.6780 80	0.02682 0.8518 51

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations					
	O2E_r6	O2E_r7	O2E_r8	O2E_r9	O2E_r10
O2E_r1 BCS: Commands % Agreement	-0.05025 0.7581 40	0.90666 <.0001 81	0.15854 0.1601 80	-0.16293 0.1487 80	-0.02093 0.8538 80
O2E_r2 BCS: Positive Attention % Agreement	-0.00922 0.9556 39	0.33374 0.0032 76	-0.00809 0.9447 76	-0.01001 0.9316 76	-0.07427 0.5237 76
O2E_r3 BCS: Negative Attention % Agreement	-0.04541 0.8293 25	0.41785 0.0139 34	0.36941 0.0315 34	-0.38351 0.0252 34	0.05500 0.7574 34
O2E_r4 BCS: Compliance % Agreement	0.12797 0.4313 40	0.66466 <.0001 81	0.17996 0.1102 80	-0.19675 0.0803 80	-0.04714 0.6780 80

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations					
	O2E_r6	O2E_r7	O2E_r8	O2E_r9	O2E_r10
O2E_r5 BCS: Noncompliance % Agreement	0.17673 0.3025 36	0.43661 0.0012 52	-0.00564 0.9687 51	-0.00498 0.9724 51	0.02682 0.8518 51
O2E_r6 BCS: Disruptive Behavior % Agreement	1.00000 40	0.06588 0.6863 40	-0.28490 0.0748 40	0.28490 0.0748 40	-0.41852 0.0072 40
O2E_r7 BCS: Total % Agreement	0.06588 0.6863 40	1.00000 81	0.15555 0.1683 80	-0.16768 0.1371 80	-0.08930 0.4309 80
O2E_r8 IRS: # of Items with Reliability of 0	-0.28490 0.0748 40	0.15555 0.1683 80	1.00000 80	-0.99288 <.0001 80	0.48154 <.0001 80
O2E_r9 IRS: Sum of All Reliability Ratings	0.28490 0.0748 40	-0.16768 0.1371 80	-0.99288 <.0001 80	1.00000 80	-0.47901 <.0001 80
O2E_r10 IRS: # of Items with Exact Agreement	-0.41852 0.0072 40	-0.08930 0.4309 80	0.48154 <.0001 80	-0.47901 <.0001 80	1.00000 80

Appendix – SAS Scoring Program
[\(See Instrument Summary\)](#)