

Western University

Scholarship@Western

---

PRECISEe Preschool Speech and Language  
Publications

PRECISE: Practice-based Research  
Collaborative in Speech-Language Pathology

---

1-12-2023

## Interrater Reliability of the FOCUS-34: Parent-to-Parent and Parent-to-Clinician

Barbara Jane Cunningham  
*Western University*, [bjcunningham@uwo.ca](mailto:bjcunningham@uwo.ca)

Peter Rosenbaum  
*McMaster University*, [rosenbau@mcmaster.ca](mailto:rosenbau@mcmaster.ca)

Anastasia Nepotiuk  
*Western University*

Nancy Thomas-Stonell  
*McMaster University*, [nthomasstonell@gmail.com](mailto:nthomasstonell@gmail.com)

Follow this and additional works at: <https://ir.lib.uwo.ca/precisepreschoolpubs>



Part of the [Speech Pathology and Audiology Commons](#)

---

### Citation of this paper:

Cunningham, Barbara Jane; Rosenbaum, Peter; Nepotiuk, Anastasia; and Thomas-Stonell, Nancy, "Interrater Reliability of the FOCUS-34: Parent-to-Parent and Parent-to-Clinician" (2023). *Communication Disorders Quarterly* 27.

<https://ir.lib.uwo.ca/precisepreschoolpubs/27>

1 **Abstract**

2 This brief report presents interrater reliability data for the Focus on the Outcomes of  
3 Communication Under Six (FOCUS-34) between parents, and between parents and speech-  
4 language pathologists (SLPs). Reliability for all three raters combined was good to excellent  
5 across three assessments. Reliability for pairs of raters was variable but generally good.

6  
7 **Introduction**

8 The Focus on the Outcomes of Communication Under Six (FOCUS) is a widely-used parent  
9 report measure that captures changes in children’s functional communication skills during  
10 speech-language therapies (Thomas-Stonell et al., 2010). Both versions (FOCUS-50 and FOCUS-  
11 34) are ideally completed by a parent, but if parents are absent, a clinician version is available.  
12 The FOCUS has demonstrated differential change over time in children with different functional  
13 abilities and communication impairments (Cunningham et al., 2021). Psychometric assessments  
14 have documented evidence of internal consistency (Thomas-Stonell et al., 2009), and  
15 convergent and discriminant validity (e.g., Thomas-Stonell, Oddson et al., 2013; Washington et  
16 al., 2013), and good interrater reliability between clinician raters (Oddson et al., 2013).  
17 Interrater reliability between parents and speech-language pathologist (SLP) raters was found  
18 to be good in one study, with fair reliability for change scores (Thomas-Stonell, Oddson et al.,  
19 2013).

20 Thomas-Stonell, Oddson et al. (2013) found good agreement between parents’ and  
21 SLPs’ ratings across multiple assessments, however, there was less agreement about change  
22 scores, and only mothers participated in the study. Additional evidence is needed to

23 understand interrater reliability between parents (including fathers) and SLPs. To date, no data  
24 have been published showing reliability between parents. This evidence is important for  
25 clinicians using the FOCUS in clinical programs (e.g., Cunningham et al., 2018) since the same  
26 parent cannot always complete all assessments. While the FOCUS-50 has been used most in  
27 research, the FOCUS-34 is used often in clinical settings (Cunningham et al., 2021), as in this  
28 study. This study explored the interrater reliability of FOCUS-34 total and change scores  
29 between parents, and between parents and SLPs. We hypothesized good agreement for both  
30 total FOCUS-34 and FOCUS-34 *change* scores. **Strength of agreement is further defined below.**

## 31 **Method**

### 32 **Ethical Approval**

33 This study was associated with a quality improvement contract involving the  
34 participating clinic, so formal ethics review was not required.

### 35 **Participants**

36 Twelve SLP volunteers from one community clinic recruited parent participants and  
37 collected data for this study. All families seen for assessment by these SLPs were invited to  
38 participate, but only families where both parents could attend the first appointment in person  
39 were included to ensure all participants had the same knowledge of the FOCUS-34. Data for 24  
40 children were reported, however complete data were available for only 13 children. Data were  
41 considered complete if both parents and the SLP submitted data for the first two assessments.

### 42 **Materials**

43 *FOCUS-34*. The FOCUS-34 is parent-report measure. Parents rate items about children's  
44 usual communicative participation on 7-point Likert scales (Oddson et al., 2019). Scores range

45 from a minimum of 34 to a maximum of 238, however it is the *change* score that is used to  
46 determine whether children have made clinically meaningful change, namely 11 or more points  
47 on the FOCUS-34 (Thomas-Stonell et al., 2020). The 11-point criterion was derived both  
48 statistically and clinically based on parents' and SLPs' judgements of whether meaningful  
49 change occurred (Oddson et al., 2019; Thomas-Stonell et al., 2013). Written instructions and  
50 definitions of communication terms are provided for those completing the FOCUS (Thomas-  
51 Stonell et al., 2020).

52 *Informal Data Collection Form.* An informal data collection form captured basic  
53 demographic and service-based information at each assessment (i.e., SLP's name, child's study  
54 ID, child's age in months, and child's level of communicative function as described using the  
55 Communication Function Classification System (CFCS), a tool for categorizing children's abilities  
56 into one of five functional levels (Hidecker et al., 2011). Together with parents, SLPs identify a  
57 child's CFCS level by considering all methods of communication and how children usually  
58 engage in everyday situations requiring communication (see [cfcs.us](http://cfcs.us) for more information).

## 59 **Procedures**

60 Participating SLPs reviewed a copy of the FOCUS-34 manual and attended an  
61 information session where administration procedures were described. The FOCUS-34 was  
62 completed independently by all participants (both caregivers and the SLP) at up to three  
63 assessment points: (1) when the child first attended therapy, (2) at the end of that therapy  
64 block (average = 4.25 months between), and (3) at a re-assessment following a period of no-  
65 treatment (average = 5.39 months from time 2). Ratings were based on participants' direct and

66 informal observations of the child in various contexts. FOCUS-34 forms were submitted to a  
67 local coordinator who submitted de-identified files to the researchers.

## 68 **Data analysis**

69 Descriptive statistics were used to profile child participants. Intraclass correlation  
70 coefficients (ICC) were calculated using a one-way random effects model (Koo & Li, 2016) to  
71 determine level of agreement for FOCUS-34 scores at each assessment and for change scores.  
72 Agreement was calculated separately for each change interval (e.g., change from Time 1-2), and  
73 within each change interval for agreement between different pairs of raters (e.g., mother-to-  
74 father agreement) where possible. With 80% power and alpha set at 0.05, a minimum sample  
75 size of 7 for the 2 × 2 ratings, and 4 for the 3 x 3 categorizations was required to detect an ICC  
76 coefficient of .80 (Bujang & Baharum, 2017). If the minimum sample size was not met, ICC  
77 coefficients were not calculated. ICC values < 0.5 were interpreted as 'poor' reliability, 0.5-0.75  
78 were 'moderate', 0.75-0.9 were 'good' and > 0.9 were 'excellent' (Koo & Li, 2016).

## 79 **Results**

80 Results are presented for 12 participants. One case was identified as an outlier (i.e.,  
81 more than 3 SD from the mean difference between raters' scores) and removed (Oddson et al.,  
82 2013). The average age of these children at the first assessment was 36.5 months ( $SD=3.9$ ).  
83 Children represented the full span of functional communication levels: CFCS Level I ( $n=1$ ), II  
84 ( $n=3$ ), III ( $n=4$ ), IV ( $n=3$ ) and V ( $n=1$ ).

85 Interrater reliability for the three raters combined was excellent at Times 1 and 2, and  
86 good at Time 3. Reliability was excellent between all pairs of raters at Time 1 and Time 2. At

87 Time 3, reliability remained good between all three raters (i.e., mothers, fathers, and SLPs), and  
88 good between fathers and SLPs, but was moderate between mothers and SLPs (see Table 1).

89 Overall interrater reliability was good when comparing change scores across the three  
90 raters for Time 1-2 (see Table 2). Reliability was good between mothers and fathers, excellent  
91 between mothers and SLPs, and good between fathers and SLPs. Good reliability was also  
92 observed between the three raters for change from Time 2-3, however the small sample size  
93 limited our ability to calculate ICCs for the pairs of raters in this change period.

#### 94 Discussion

95 Several clinical and research programs have adopted the FOCUS as an outcome measure  
96 and use it to assess the impact of interventions for children across impairment types and levels  
97 of ability (Cunningham et al., 2021). To ensure research rigor, the same caregiver would  
98 complete all assessments. In practice this is not always feasible, so it was important to  
99 determine whether the FOCUS-34 could be completed reliably by different raters. Results  
100 indicated good interrater reliability between three raters (i.e., mothers, fathers and SLPs) across  
101 multiple assessment points. Reliability was lower at Time 3 for mother-SLP ratings. This may be  
102 due to the smaller sample size ( $N=7$ ), combined with one large disagreement between two  
103 raters.

104 Across assessments most mothers scored their children higher than the SLPs. This could  
105 be due to the vastly greater opportunities mothers have to observe their children's  
106 communication in multiple contexts, or because familiarity facilitates communication with their  
107 child, or perhaps because children perform differently in different environments. It is also

108 possible that SLPs see more children communicate and may unknowingly compare between  
109 children (not the intended use of the FOCUS), leading to their lower ratings.

110 Outcome measures are designed to assess change. Therefore, it was essential to  
111 evaluate interrater reliability for change scores. Reliability for change scores in this study was  
112 good for both change intervals. When examining change between Times 1 and 2, moderate  
113 reliability was noted between the father-SLP pairs. While we did not collect this information,  
114 we suspect, based on past research with this program (e.g., Thomas-Stonell, Washington et al.  
115 2013), that it was primarily mothers who brought the children to speech-language therapy  
116 sessions. Therefore, the mothers may have had more ongoing contact with the SLPs and more  
117 opportunities to discuss their child's progress. **It is also possible that mothers had more**  
118 **opportunity to learn about their child's communication attempts through observation at home**  
119 **and during therapy sessions. These interpretations** may also explain the slightly higher  
120 reliability for mothers when compared to fathers. Ongoing contact between parent and SLP  
121 may be an important factor in ensuring high reliability over time. It is thus recommended that a  
122 consistent rater complete the FOCUS-34 over time, but our results indicate that, if needed,  
123 different raters can reliably complete the FOCUS-34. **It was beyond the scope of this study to**  
124 **explore differences in participants' ratings, but future work should explore whether there are**  
125 **differences in how parents and SLPs perceive specific aspects of communication.**

126 It is important to note that larger sample sizes are typically used in assessing reliability  
127 (Koo & Li, 2016). This study used a small sample of convenience obtained by SLPs recruiting  
128 parent volunteers. As such, selection bias is possible, **and the predictive value of our results**  
129 **may be limited.** The FOCUS-34 was designed to be applicable for children with a variety of

130 communication differences, and while the children in this study represent a range of functional  
131 abilities, they may not be representative of all preschoolers attending speech-language therapy.

132 **Conclusions**

133 While it remains preferable to use a consistent rater to complete the FOCUS-34 over  
134 time, results demonstrate that different raters can reliably complete it. This finding has  
135 practical implications in large clinical programs. Results also suggest that reliability may be  
136 improved with ongoing contact between parents and SLPs. Future research should include  
137 replication of this study with a larger sample size and evaluate reliability of the FOCUS-34 when  
138 completed by teachers and early childhood educators – with whom SLPs often consult.



139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159

## References

Bujang, M. A. & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Sciences*, 12(1), 1-11.

Cunningham, B. J., Hanna, S. E., Rosenbaum, P., Thomas-Stonell, N. & Oddson, B. (2018). Factors contributing to preschoolers' communicative participation outcomes: Findings from a population-based longitudinal cohort study in Ontario Canada. *American Journal of Speech-Language Pathology*, 27(2), 737-750.

Cunningham, B. J., Thomas-Stonell, N., & Rosenbaum, P. (2021). Assessing communicative participation in preschool children with the Focus on the Outcomes of Communication Under Six: a scoping review. *Developmental Medicine & Child Neurology*, 63: 47-53.

Hidecker, M. J., Paneth, N., Rosenbaum, P.L., et al. (2011) Developing and validating the Communication Function Classification System for individuals with cerebral palsy. *Developmental Medicine and Child Neurology*, 53, 704–10.

Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163.

Oddson, B., Thomas-Stonell, N., Robertson, B., & Rosenbaum, P. (2019). Validity of a streamlined version of the Focus on the Outcomes of Communication Under Six: Process and outcome. *Child: Care, Health and Development*, 45(4), 600-605.

Oddson, B., Washington, K., Robertson, B., Thomas-Stonell, N., & Rosenbaum, P. (2013). Inter-rater Reliability of Clinicians' Ratings of Preschool Children Using the FOCUS©: Focus on

160 the Outcomes of Communication Under Six. *Canadian Journal of Speech-Language*  
161 *Pathology and Audiology*, 37(2), 170- 174.

162 Thomas-Stonell, N., Oddson, B., Robertson, B., & Rosenbaum, P. (2009). Predicted and  
163 observed outcomes in preschool children following speech and language treatment:  
164 Parent and clinician perspectives. *Journal of Communication Disorders*, 42(1): 29-42.

165 Thomas-Stonell, N. L., Oddson, B., Robertson, B., & Rosenbaum, P. L. (2010). Development of  
166 the FOCUS© (Focus on the Outcomes of Communication Under Six), a communication  
167 outcome measure for preschool children. *Developmental Medicine and Child Neurology*,  
168 52(1), 47–53.

169 Thomas-Stonell, N., Oddson, B., Robertson, B., & Rosenbaum, P. (2013). Validation of the  
170 Focus on the Outcomes of Communication under Six outcome measure. *Developmental*  
171 *Medicine & Child Neurology*, 55(6), 546-552.

172 Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B. & Rosenbaum, P. (2013).  
173 Measuring communicative participation using the FOCUS©: Focus on the Outcomes of  
174 Communication Under Six. *Child: care, health and development*, 39(4), 474-480.

175 Thomas-Stonell, N., Cunningham, B. J., Robertson, B., Rosenbaum, P. (2020). *The Focus on the*  
176 *Outcomes of Communication Under Six (FOCUS©) Manual, 2<sup>nd</sup> Edition*. CanChild.

177 Washington, K., Thomas-Stonell, N., Oddson, B., McLeod, S., Warr-Leeper, G., Robertson, B. &  
178 Rosenbaum, P. (2013). Construct validity of the FOCUS©: A communicative participation  
179 outcome measure for preschool children. *Child: Care, Health and Development*, 39(4),  
180 481-489.

181 **Table 1.** Interrater reliability for total FOCUS scores at each assessment point  
 182  
 183

Reliability pairs	Time 1 ( <i>n</i> = 12)	Time 2 ( <i>n</i> = 12)	Time 3 ( <i>n</i> = 7)
<b>All 3 raters combined</b>	<b>ICC = 0.97 (0.93-0.99)</b> <b>F = 37.15, <i>p</i> &lt; .01</b> <b>Int = Excellent</b>	<b>ICC = 0.95 (0.88-0.99)</b> <b>F = 21.64, <i>p</i> &lt; .01</b> <b>Int = Excellent</b>	<b>ICC = 0.87 (0.55-0.98)</b> <b>F = 7.73, <i>p</i> &lt; .01</b> <b>Int = Good</b>
Mother/ Father	ICC = 0.96 (0.86-0.99) F = 24.39, <i>p</i> < .01 Int = Excellent	ICC = 0.95 (0.84-0.99) F = 20.87, <i>p</i> < .01 Int = Excellent	ICC = 0.86 (0.29-0.98) F = 7.23, <i>p</i> = .01 Int = Good
Mother/ SLP	ICC = 0.97 (0.91-0.99) F = 36.30, <i>p</i> < .01 Int = Excellent	ICC = 0.93 (0.78-0.98) F = 15.06, <i>p</i> < 0.1 Int = Excellent	ICC = 0.65 (-0.80-0.94) F = 2.84, <i>p</i> = .10 Int = Moderate
Father/ SLP	ICC = 0.95 (0.83-0.98) F = 19.22, <i>p</i> < .01 Int = Excellent	ICC = 0.91 (0.70-0.97) F = 11.00, <i>p</i> < .01 Int = Excellent	ICC = 0.87 (0.35-0.98) F = 7.92, <i>p</i> < .01 Int = Good

184  
 185 *n*: number of children, *ICC*: Intraclass correlation coefficient, *Int*: reliability interpretation, *SLP*:  
 186 Speech-Language Pathologist  
 187  
 188 *Note*: According to Koo and Li (2016), ICC values of less than 0.5 = ‘poor’ reliability, values  
 189 between 0.5 to 0.75 = ‘moderate’ reliability, values between 0.75 to 0.9 = ‘good’ reliability and  
 190 greater than 0.9 were ‘excellent’ reliability

191 **Table 2.** Interrater reliability for FOCUS change scores  
 192  
 193

Reliability pairs	Time 1 to Time 2 ( <i>n</i> = 12)	Time 2 to Time 3 ( <i>n</i> = 7)
<b>All 3 raters combined</b>	<b>ICC = 0.84 (0.59-0.95)</b> <b>F = 6.39, <i>p</i> &lt; .01</b> <b>Int = Good</b>	<b>ICC = 0.82 (0.37-0.97)</b> <b>F = 5.56, <i>p</i> = .004</b> <b>Int = Good</b>
Mother/Father	ICC = 0.77 (0.24-0.93) F = 4.37, <i>p</i> = .009 Int = Good	NA
Mother/SLP	ICC = 0.90 (0.65-0.97) F = 9.56, <i>p</i> < .001 Int = Good	NA
Father/SLP	ICC = 0.60 (-0.32-0.88) F = 2.52, <i>p</i> = .06 Int = Moderate	NA

194 *n*: number of children, *ICC*: Intraclass correlation coefficient, *Int*: reliability interpretation *SLP*:  
 195 Speech-Language Pathologist, *NA*: not applicable

196  
 197 *Note*: According to Koo and Li (2016), *ICC* values of less than 0.5 = 'poor' reliability, values  
 198 between 0.5 to 0.75 = 'moderate' reliability, values between 0.75 to 0.90 = 'good' reliability and  
 199 greater than 0.90 were 'excellent' reliability

200  
 201