

## Study Purpose

In this study, we evaluated interrater agreement (IRA) of the Quality Standards Assessment (QSA). IRA is the absolute consensus in rating scores from multiple raters on the same targets. High IRA justifies aggregation of scores from multiple raters.



- **Measure of IRA:** We used the  $r_{WG(J)}$  index as the measure of Interrater Agreement (IRA), which defines agreement in terms of the proportional reduction in error variance (LeBreton & Senter, 2008). The  $r_{WG(J)}$  index is the most popular measure of IRA, and measures IRA by comparing the observed variance in ratings furnished by multiple judges of a single target to the variance one would expect when the judges responded randomly.

Formula to calculate of  $r_{WG(J)}$  is:

$$r_{WG(J)} = \frac{J \left( 1 - \frac{\bar{S}_{Xj}^2}{\sigma_E^2} \right)}{J \left( 1 - \frac{\bar{S}_{Xj}^2}{\sigma_E^2} \right) + \left( \frac{\bar{S}_{Xj}^2}{\sigma_E^2} \right)}$$

- J indicates the number of essentially parallel items.
- $\bar{S}_{Xj}^2$  Indicates the mean of the observed variance for J essentially parallel items.
- $\sigma_E^2$  Indicates the expected variance obtained from a theoretical null distribution representing a complete lack of agreement among judges.
- We used the  $r_{WG(J)}$  value of 0.50, the cut point of moderate agreement to indicate acceptable IRA, since most group homes were rated by only 2 raters of the same title, which is a small number of raters and can attenuate the value of  $r_{WG(J)}$ . In addition, we used the  $r_{WG(J)}$  value of 0.70, the cut point of strong agreement to indicate strong IRA.

## Study Sample

We used state-wide data collected in 2018-2019 from 189 group homes (GHs) licensed by the Florida State Department of Children and Families (DCF).

Among the 189 GHs, 142 of them were rated by 2 or more youth with the youth form of the Florida QSA; 25 GHs were rated by 2 or more external raters from case management agencies with the contract agency form; 131 were rated by 2 or more direct care workers within the GH with the direct care worker form; 71 were rated by 2 or more directors or supervisors within the GH with the director form; 31 were rated by 2 or more licensure specialists from Florida DCF with the licensing form.

## Study Results

The results showed that for the forms used by direct care workers, director/supervisors, and lead agency staff, Domains 2-7 showed acceptable IRAs, indicated by that at least 60% GHs have moderate-high IRA with the values of  $r_{WG(J)}$  ranging from 0.50 and 1. For the youth form, Domains 2 and 6 showed satisfactory IRAs, indicated by that at least 60% GHs have moderate-high IRA. For the other domains on the youth form, at least 50% GHs have moderate-high IRA. For the licensing form, all the domains showed acceptable IRAs.

### Percentages of GHs reaching moderate agree ( $r_{WG(J)}$ greater than 0.50)

	Contract agency form	Direct care worker form	Director form	Youth Form	Licensing Form
domain 1 (14 item)	70	50	40	50	100
domain 2 (12 items)	80	60	80	60	100
domain 3 (8 items)	80	70	90	50	100
domain 4 (9 items)	80	70	80	50	97
domain 5 (7 items)	60	80	80		94
domain 6 (18 items)	90	70	80	60	97
domain 7 (9 items)	90	80	90	50	94
domain 8 (7 items)	50	70	50	50	85

### Percentages of GHs reaching strong agree ( $r_{WG(J)}$ greater than 0.70)

	Contract agency form	Direct care worker form	Director form	Youth Form	Licensing Form
domain 1 (14 item)	60	50	40	50	94
domain 2 (12 items)	70	60	80	40	97
domain 3 (8 items)	80	60	90	60	100
domain 4 (9 items)	80	70	80	50	97
domain 5 (7 items)	60	70	80		90
domain 6 (18 items)	90	70	80	50	90
domain 7 (9 items)	80	70	90	50	94
domain 8 (7 items)	50	60	50	40	81

## Discussion

- Among the five forms of QSA, most domains showed acceptable IRA. Therefore, these finding lend support for aggregating ratings from multiple raters of the same title to provide a composite score on quality of residential care. The strong IRA between licensure specialists lends support for having one licensure specialist to rate each GH.
- Future research needs to examine the inter-rater agreements between raters of different titles.
- Our finding further supports that QSA has the potential to be used as a reliable measure to measure quality of group care, in response to the Family First Prevention Services Act (Family First), which limits the use of group care to Qualified Residential Treatment Programs.