Intraclass Correlation Coefficient

The intraclass correlation coefficient (ICC) is a general statistic that is used in multilevel modeling, ANOVA, psychometrics, and other areas. It is a measure of the degree of clustering within groups (or classes), but it also represents a complementary concept, the degree of variability between groups. If we consider that the variance of the dependent variable, Y_{ij}, can be partitioned into variance due to individual variation within a group (e.g., math achievement among students within a school) and variation across groups (e.g., average math achievement among schools), then we can create a ratio of the amount of variance due to groups (τ_0^2) relative to the total variance of Y_{ij} ($\sigma^2 + \tau_0^2$).

$$ICC = \rho = \frac{\tau_0^2}{\sigma^2 + \tau_0^2}$$

The resulting value is theoretically falls between 0 and 1.0 because it is a proportion of variance,[1] with higher values reflecting greater between-group variability. The ICC is therefore giving similar information to the test of the between-group variance. Although one can compute a standard error, significance test, and confidence intervals for the ICC, I consider this to be generally irrelevant, because a trivial ICC will result in biased standard errors (and thus Type I errors) if clustering is ignored (Barcikowski, 1981; Huang, 2018), and the significance test of the variance of the intercept provides information about significant group variability. The above formula is commonly used for computed ICC, but others have been proposed (Bliese, 1988; Shrout & Fleiss, 1979).[2] Nonetheless, ICC is a valuable descriptive statistic used as a preliminary step for better understanding the proportion of variance due to group differences.

Homogeneity Within Groups and Heterogeneity Between Groups

To illustrate the range of the ICC, consider a hypothetical example involving the number of client contacts, nested within five therapists. In Example 1, all variance is between groups, and in Example 2, all variance is within groups.

Example 1			
	i	j	Y_{ij}
$\bar{Y}_j = 6$, $\sigma_j^2 = 0$	1	1	6
	2	1	6
	3	1	6
	4	1	6
	5	1	6
$\bar{Y}_j = 7$, $\sigma_j^2 = 0$	6	2	7
	7	2	7
	8	2	7
	9	2	7
	10	2	7
$\bar{Y}_j = 8$, $\sigma_j^2 = 0$	11	3	8
	12	3	8
	13	3	8
	14	3	8
	15	3	8
$\bar{Y}_j = 9$, $\sigma_j^2 = 0$	16	4	9
	17	4	9
	18	4	9
	19	4	9
	20	4	9
$\bar{Y}_j = 10$, $\sigma_j^2 = 0$	21	5	10
	22	5	10
	23	5	10
	24	5	10
	25	5	10
	$\bar{Y}_{..} = 8$, $\sigma^2_{total} = 2.083$ $\rho = 1.0$		

Example 2			
i	j	Y_{ij}	
1	1	6	$\bar{Y}_j = 8$, $\sigma_j^2 = 2.5$
2	1	7	
3	1	8	
4	1	9	
5	1	10	
6	2	6	$\bar{Y}_j = 8$, $\sigma_j^2 = 2.5$
7	2	7	
8	2	8	
9	2	9	
10	2	10	
11	3	6	$\bar{Y}_j = 8$, $\sigma_j^2 = 2.5$
12	3	7	
13	3	8	
14	3	9	
15	3	10	
16	4	6	$\bar{Y}_j = 8$, $\sigma_j^2 = 2.5$
17	4	7	
18	4	8	
19	4	9	
20	4	10	
21	5	6	$\bar{Y}_j = 8$, $\sigma_j^2 = 2.5$
22	5	7	
23	5	8	
24	5	9	
25	5	10	
$\bar{Y}_{..} = 8$, $\sigma^2_{total} = 2.083$ $\rho = 0$			

---

[1] Because the estimate of between-group variance τ_0^2, written as $\hat{\tau}^2$ in the discussion of ICC in your text, can be negative, the ICC might be negative. This is expected to be a rare occurrence when the between-group variance is 0 or near zero. See Snijders and Bosker (2012) p. 21 for more detail.

[2] ICC in the above equation is the same as ICC1 described by Bliese (1988) and ICC(1,1) described by Shrout and Fleiss (1979). See also Shieh (2016).

Computation of ICC using Intercept-Only (Empty) Model from HSB Example
SPSS Output

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		39.14832	.6606447	59.258	.000	37.8746616	40.4648133
Intercept [subject	Variance	8.6140248	1.0788036	7.985	.000	6.7391217	11.0105479

a. Dependent Variable: mathach.

R Output

```
> #nlme provides standard deviations of the random effects by default, use VarCorr to obtain variances
> VarCorr(model)
schoolid = pdLogChol(1)
            Variance  StdDev
(Intercept) 8.614025 2.934966
Residual   39.148322 6.256862
```

In, R, ICC can be computed using the ICC1.lme function that is in the `psychometric` package.[3]

```
> library(psychometric)
 > ICC1.lme(mathach,schoolid,data=mydata)
[1] 0.1803518
```

HLM Output

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	2.93501	8.61431	159	1660.23259	<0.001
level-1, r	6.25686	39.14831			

Statistics for current covariance components model
Deviance = 47116.793477
Number of estimated parameters = 2

Each of the software programs gives nearly identical estimates for the between group variance (8.614) and the within-group variance (39.148).

$$\rho = \frac{\tau_0^2}{\sigma^2 + \tau_0^2} = \frac{8.614}{39.148 + 8.614} = .18$$

The proportion of the total variance of Y_{ij} that is due to schools is .18 or 18%.

Unconditional vs. Conditional ICC

The above example uses the “unconditional” ICC, which is the most commonly used in practice, but the ICC can also be computed from the same components taken from an analysis with one or more predictors, and, thus, represents the “conditional” ICC—between group variability after controlling for the predictors.

References

Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational and Behavioral Statistics*, *6*(3), 267-285.

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods, 1*(4), 355-373.

Huang, F. L. (2018). Multilevel modeling and ordinary least squares regression: How comparable are they?. *The Journal of Experimental Education, 86*(2), 265-281.

Shieh, G. (2016). Choosing the best index for the average score intraclass correlation coefficient. *Behavior Research Methods, 48*, 994-1003.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

[3] This is the simplest approach here but there are a number of other functions in R for obtaining ICC, including when using the `lme4` and `performance` package:

```
> nullmodel = lmer(mathach ~ (1|schoolid), data=mydata)
> icc(nullmodel)
```