

**Unbalanced and Missing Data Generalizability Analyses:
A tutorial based on Henderson (1953), Brennan (2001a)**

By: Tyler J. Smith and Theresa J.B. Kline 15 January 2025

This tutorial assumes a fairly good working knowledge of generalizability theory and the terms used in its analysis. If the terms “facet”, “variance component”, “generalizability”, “dependability”, “G-Study” and “D-study” are unfamiliar, it is strongly advised that the reader review an introduction to generalizability procedures using balanced/non-missing data. There are many choices for this such as Briesch, Swaminathan, Welsh, and Chafouleas (2014).

While data sets with balanced facets and with no missing data points are relatively straightforward to analyze for generalizability, the same is not true for data sets not meeting these criteria. In fact many data sets (including those used in machine learning) that researchers wish to subject to generalizability analyses are either: 1) unbalanced (have different numbers of cases in each combination/nesting of variables); and/or 2) are missing data. There have been few freely available options to analyze such data – exceptions include the urGENOVA (Brennan, 2001b) and G-String_V (Bloch & Norman, 2012; 2023) programs. The urGENOVA program allows for both unbalanced data sets and missing data but has an esoteric user interface and does not calculate generalizability coefficients as part of the output. The G-String_V program has an intuitive user interface and produces generalizability and decision-study outputs but does not allow for missing data.

Generalizability analyses are dependent on **variance components** for their calculation. In balanced designs with no missing data, these are readily calculated using the Sums of Squares/Mean Squares generated in typical ANOVA analyses. When these criteria are not met, another approach (“analogous ANOVA”) has been used (Brennan, 2001a) and is based on the “Method 1” approach introduced by Henderson (1953) to generate variance components in data sets that are unbalanced and/or have missing data. It is mathematically simple - using cell frequency counts and squared/summed values of the data point values to calculate the variance components. The generalizability/dependability calculations can follow from there.

Henderson 1953

The first step in this document is to “walk” the reader through the “Method 1” approach to calculate **variance components** introduced by Henderson (1953) and the one utilized by Brennan (2001a) in creating the urGENOVA generalizability program. Note that calculating generalizability coefficients is NOT the end goal of Henderson - it is just to generate variance coefficients, which indicate how the variance in a set of data is distributed among its various components (or “facets” in the vernacular of generalizability theory). It will be simplest to show how these work with an actual data set.

Table 1 below shows the data from Henderson’s (1953) Table 1. Its cells contain the butterfat levels at first lactations of 4 cow herds inseminated by 3 different sires over 4 years. The entries in each cell are somewhat unique – they represent number of cows in each herd (first number) and butterfat content (second number) in the first lactation of the year. The entry “3 – 1414” represents the total butterfat (1414) summed over 3 different cows. Thus, cows are nested within Herd/Sire combinations (C: HS).

Table 1. Henderson’s butterfat data set “Year(A) X Herd (H) X Sire (S)”

| HERD | SIRE | YEAR | | | | TOTAL |
|-------|------|----------|----------|-----------|-----------|----------|
| | | 1 | 2 | 3 | 4 | |
| 1 | 1 | 3 - 1414 | 2 – 981 | | | 5 – 2395 |
| 1 | 2 | | 4 – 1766 | 2 – 862 | | 6 – 2628 |
| 1 | 3 | | | | 5 - 1609 | 5 – 1609 |
| 2 | 1 | 1 - 405 | 3 – 1270 | | | 4 – 1674 |
| 2 | 2 | | | 5 – 2109 | | 5 – 2109 |
| 2 | 3 | | | 4 – 1563 | 2 - 740 | 6 - 2303 |
| 3 | 1 | | 3 – 1705 | | | 3 – 1705 |
| 3 | 2 | | 4 – 2310 | 2 – 1134 | | 6 – 3444 |
| 4 | 1 | 3 - 1113 | 5 - 1951 | | | 8 – 3064 |
| 4 | 3 | | | 3 - 1291 | 6 - 2457 | 9 – 3948 |
| TOTAL | | 7 - 2931 | 21-9983 | 16 - 6959 | 13 - 4806 | 57-24679 |

It is worth noting:

1. Most of the data are missing.
2. The 'cows' variable is not of particular interest in the study. If it was, the actual butterfat values for each of the cows within the HS combinations would be listed. Thus, the focus of the rest of the analysis is on the crossed (but not fully crossed) Year(A) X Herd (H) X Sire (S). However, this issue DOES come up later in this tutorial. For now, we will set it aside.
3. That for Herds 3 and 4, only 2 sires are crossed with them; sires 1 and 2 for Herd 3 and sires 1 and 3 for Herd 4. This means that there are 10 HS combinations. It is as though there are "missing data" for sire 3 herd 3, and sire 2 herd 4.
4. Henderson writes the equation that represents the linear model of the data is as follows: Note that there is a 'k' in the error term – this term includes random error PLUS also the cows:HS nested term (which is not assessed or accounted for in the model).

$$Y(hijk) = \mu + A(\text{Year})h + H(\text{Herd})i + S(\text{Sire})j + HS(\text{Herd} \times \text{Sire})ij + \text{Error}(hijk).$$

Assumptions:

1. All variables are random
2. All effects (except the grand mean) are uncorrelated with each other, have means of 0 and have variances.

Step 1: Obtain the "T" values (uncorrected sums of squares) for each facet by summing over the squared values for each level of the facet and divide by each level's relevant sample size ("^" indicates exponential). Henderson uses summation notation in his article – however actual values will make the computations more concrete.

$$A(\text{year}) = \text{Sum } (2931^2)/7 + (9983^2)/21 + (6959^2)/16 + (4806^2)/13 = 10,776,451$$

$$H(\text{herd}) = \text{Sum } ((2395+2628+1609)^2)/(5+6+5) + ((1674+2109+2303)^2)/(4+5+6) + ((1705+3444)^2)/(3+6) + ((3064+3748)^2)/(8+9) = 10,893,666$$

$$S(\text{sire}) = \text{Sum } ((2395+1674+1705+3064)^2)/(5+4+3+8) + ((2628+2109+3444)^2)/(6+5+6) + ((1609+2303+3748)^2)/(5+6+9) = 10,776,278$$

$$HS(\text{herd} \times \text{sire}) = \text{Sum } (2395^2)/5 + (2628^2)/6 + (1609^2)/5 + (1674^2)/4 + (2109^2)/5 + (2303^2)/6 + (1705^2)/3 + (3444^2)/6 + (3064^2)/8 + (3748^2)/9 = 10,970,369$$

$$\text{Mean (Henderson uses "CF" for this term)} = \text{Sum over all values } (24679^2)/57 = 10,685,141$$

$$\text{Total} = 11,124,007^*$$

Total = **11,124,007*** This would normally be accomplished by squaring and summing **each individual cell** value (all 57 of them). In this data set this would mean we would have to have the individual butterfat values for EACH "cow". Instead, Henderson provides the sum of the butterfat values across cows in the herd. For example, in the cell 3-1414, this means that 3 cows provided 1414 butterfat. We DO NOT know how the 1414 butterfat is distributed between the 3 cows. If we did, we would be able to calculate the nested effect of cows within herds (c:H), that we are not provided. Henderson later in the paper does provide a T-value for the total (Table 6, p. 233) of **11,124,007**. I tried generating this value a couple of different ways:

1. Taking the overall total 24679 and dividing by 57 (433). Then squaring this value 57 times and summing. This gave me a T-value of 10,686,873.
2. Taking each herd value, dividing it by the number of cows in the herd, then "counting" that value for each cow for that herd. This was repeated 57 times. Then these values were squared then summed. This gave me a T-value of 10,973,517.
3. Taking each herd value, dividing it by the number of cows in the herd. Then averaging across those values (440). Then squaring this value 57 times and summing. This gave me a T-value of 11,035,200.

Clearly the actual values for each cow would need to be provided to get the actual total T value for this analysis.

Step 2: Obtain the estimated coefficients by which the variance components will be multiplied

So far, things have been straightforward. The next step is more complicated and tedious. The coefficients of μ^2 , $\sigma^2(\text{year})$, $\sigma^2(\text{herd})$, $\sigma^2(\text{sire})$, $\sigma^2(\text{herd} \times \text{sire})$, $\sigma^2(\text{error})$ are *estimated* using the SAMPLE SIZES on which they are based that contribute to the calculated T-values. This is tedious because the data set is unbalanced and there are lots of missing data! Fortunately, Henderson provides a slow and steady summation routine that allows for the estimations. Many of the cells can be filled using the simple overall sample size of the data set (N) (these are described below). The others (noted with a “?” in Table 2) have to be estimated individually. Again, Henderson provides the summation notation for calculating these terms for those interested in seeing them.

1. For the μ^2 term, ALL effects are based on the total number of data points involved in the study (denoted N).
2. For the Total effect, all variances are also based on the total number of data points involved in the study (N).
3. For each effect (except the Mean), the variance for that effect is also for the based on the total number of data points involved in the study (N). For example, the variance of $\sigma^2(\text{year})$ for effect A (year) is N.
4. For the $\sigma^2(\text{error})$ term column, the facet sample sizes are equal to the number of levels of that effect (Year = 4, Herd = 4, Sire = 3, HS = 10 (there are 10 different HS combinations), for the Mean it is always = 1, and for the Total it is always = N.
5. This leaves 16 “non-easy” ones that require calculation and are noted as such.

Table 2. Hendersons T-values (uncorrected Sums of Squares) for each facet and coefficient sample size bases

| Facet | T-Value | μ^2 | $\sigma^2(\text{Year})$ | $\sigma^2(\text{Herd})$ | $\sigma^2(\text{Sire})$ | $\sigma^2(\text{HS})$ | $\sigma^2(\text{error})$ |
|-------------------------|--------------------|---------|-------------------------|-------------------------|-------------------------|-----------------------|--------------------------|
| A (year) | 10,776,451 | N | N | ? - 1 | ? - 2 | ? - 3 | Levels of A (4) |
| H (herd) | 10,893,666 | N | ? - 4 | N | ? - 5 | ? - 6 | Levels of H (4) |
| S (sire) | 10,776,278 | N | ? - 7 | ? - 8 | N | ? - 9 | Levels of S (3) |
| HS (herd \times sire) | 10,970,369 | N | ? - 10 | ? - 11 | ? - 12 | N | #Combinations of HS (10) |
| Mean | 10,685,141 | N | ? - 13 | ? - 14 | ? - 15 | ? - 16 | 1 |
| Total | 11,124,007* | N | N | N | N | N | N |

*See note above regarding where this value comes from

Let's begin to “fill in” the 16 different “?” terms with “brute force” calculations....

1. The $\sigma^2(\text{Herd})$ term on Year (data are collapsed across Sires). Sample size starts by calculating the squared sum of: # cows that provided data in each herd within each year, divided by the total number of cows for that year. Then sum the quotients.

Year 1: $((3^2) + (1^2) + (0^2) + (3^2))/7 +$
Year 2: $((2+4)^2) + (3^2) + ((3+4)^2) + (5^2))/21 +$
Year 3: $((2^2) + ((5+4)^2) + (2^2) + (3^2))/16 +$
Year 4: $((5^2) + (2^2) + (0^2) + (6^2))/13$
=19.51

2. The $\sigma^2(\text{Sire})$ term on Year (data are collapsed across Herds). Sample size starts by calculating the squared sum of: # cows that provided data for each sire within each year, divided by the total number of cows for that year. Then sum the quotients.

Year 1: $((3+1+0+3)^2) + ((0+0+0+0)^2) + ((0+0+0+0)^2))/7 +$
Year 2: $((2+3+3+5)^2) + ((4+0+4+0)^2) + ((0+0+0+0)^2))/21 +$
Year 3: $((0+0+0+0)^2) + ((2+5+2+0)^2) + ((0+4+0+3)^2))/16 +$
Year 4: $((0+0+0+0)^2) + ((0+0+0+0)^2) + ((5+2+0+6)^2))/13$
=39.22

3. The $\sigma^2(\text{HS})$ term on Year. Sample size starts by calculating the squared sum of:
cows that provided data for each of the 10 herd/sire combinations within each year, divided by the total number of cows for that year. Then sum the quotients.

$$\begin{aligned} \text{Year 1: } & ((3^2) + (0^2) + (0^2) + (1^2) + (0^2) + (0^2) + (0^2) + (0^2) + (3^2) + (0^2))/7 + \\ \text{Year 2: } & ((2^2) + (4^2) + (0^2) + (3^2) + (0^2) + (0^2) + (3^2) + (4^2) + (5^2) + (0^2))/21 + \\ \text{Year 3: } & ((0^2) + (2^2) + (0^2) + (0^2) + (5^2) + (4^2) + (0^2) + (2^2) + (0^2) + (3^2))/16 + \\ \text{Year 4: } & ((0^2) + (0^2) + (5^2) + (0^2) + (0^2) + (2^2) + (0^2) + (0^2) + (0^2) + (6^2))/13 \\ & = 15.10 \end{aligned}$$

4. The $\sigma^2(\text{Year})$ term on Herd (data are collapsed across Sires). Sample size starts by calculating the squared sum of:
cows that provided data in each year within each herd, divided by the total number of cows for that herd. Then sum the quotients.

$$\begin{aligned} \text{Herd 1: } & ((3^2) + ((2+4)^2) + (2^2) + (5^2))/16 + \\ \text{Herd 2: } & ((1^2) + (3^2) + ((5+4)^2) + (2^2))/15 + \\ \text{Herd 3: } & ((0^2) + ((3+4)^2) + (2^2) + (0^2))/9 + \\ \text{Herd 4: } & ((3^2) + (5^2) + (3^2) + (6^2))/17 \\ & = 21.49 \end{aligned}$$

5. The $\sigma^2(\text{Sire})$ term on Herd (data are collapsed across Years). Sample size starts by calculating the squared sum of:
cows that provided data in each herd, divided by the total number of cows for that herd. Then sum the quotients.

$$\begin{aligned} \text{Herd 1: } & (((3+2)^2) + ((4+2)^2) + (5^2))/16 + \\ \text{Herd 2: } & ((3+1)^2) + (5^2) + ((4+2)^2))/15 + \\ \text{Herd 3: } & ((3^2) + ((4+2)^2) + (0^2))/9 + \\ \text{Herd 4: } & (((3+5)^2) + (0^2) + ((3+6)^2))/17 \\ & = 24.04 \end{aligned}$$

6. The $\sigma^2(\text{HS})$ term on Herd the sample size starts by calculating the squared sum of:
cows that provided data in each herd, divided by the total number of cows for that herd. Then sum the quotients.

Note that this is the SAME value as that for the $\sigma^2(\text{Sire})$ effect on Herd. The difference is in *how* the numbers were arrived at. For the $\sigma^2(\text{Sire})$ effect on Herd, we had to sum across the Years to get the number of cows. For the $\sigma^2(\text{HS})$ effect we just have used the sum of cows across years (like having rolled the Years into a single column).

$$\begin{aligned} \text{Herd 1: } & ((5^2) + (6^2) + (5^2))/16 + \\ \text{Herd 2: } & ((4^2) + (5^2) + (6^2))/15 + \\ \text{Herd 3: } & ((3^2) + (6^2) + (0^2))/9 + \\ \text{Herd 4: } & ((8^2) + (0^2) + (9^2))/17 \\ & = 24.04 \end{aligned}$$

7. The $\sigma^2(\text{Year})$ term on Sire (data are collapsed across Herds). Sample size starts by calculating the squared sum of:
cows that provided data in each year within each sire, divided by the total number of cows for that sire. Then sum the quotients.

$$\begin{aligned} \text{Sire 1: } & ((3+1+3)^2) + ((2+3+3+5)^2) + (0^2)/(0))/20 + \\ \text{Sire 2: } & ((0^2) + ((4+4)^2) + ((2+5+2)^2) + ((0^2))/17 + \\ \text{Sire 3: } & ((0^2) + (0^2) + ((4+3)^2) + ((5+2+6)^2))/20 + \\ & = 30.33 \end{aligned}$$

8. The $\sigma^2(\text{Herd})$ term on Sire (data are collapsed across Years). Sample size starts by calculating the squared sum of:
cows that provided data in each herd within each sire, divided by the total number of cows for that sire. Then sum the quotients.

$$\begin{aligned}\text{Sire 1: } & ((3+2)^2) + ((1+3)^2) + (3^2) + ((3+5)^2) / 20 + \\ \text{Sire 2: } & ((4+2)^2) + (5^2) + ((4+2)^2) + (0^2) / 17 + \\ \text{Sire 3: } & (5^2) + ((4+2)^2) + (0^2) + ((6+3)^2) / 20 + \\ & = 18.51\end{aligned}$$

9. For the $\sigma^2(\text{HS})$ term on Sire the sample size starts by calculating the squared sum of:
cows that provided data in each sire, divided by the total number of cows for that sire. Then these quotients are summed.

Note that this is the SAME value as that for the $\sigma^2(\text{Herd})$ effect on Sire. The difference is in how the numbers were arrived at. For the $\sigma^2(\text{Herd})$ effect on Sire, we had to sum across the Years to get the number of cows. For the $\sigma^2(\text{HS})$ effect we just have used the sum of cows across years (like having rolled the Years into a single column).

$$\begin{aligned}\text{Sire 1: } & (5^2) + (4^2) + (3^2) + (8^2) / 20 + \\ \text{Sire 2: } & (6^2) + (5^2) + (6^2) + (0^2) / 17 + \\ \text{Sire 3: } & (5^2) + (6^2) + (0^2) + (9^2) / 20 + \\ & = 18.51\end{aligned}$$

10. The $\sigma^2(\text{Year})$ term on HS the sample size starts by calculating the squared sum of:
cows that provided data across years for each Herd/Sire combination, divided by the total number of cows for that Herd/Sire combination. Then sum the quotients.

$$\begin{aligned}\text{Herd/Sire 1: } & ((3^2) + (2^2) + (0^2) + (0^2)) / 5 + \\ \text{Herd/Sire 2: } & ((0^2) + (4^2) + (2^2) + (0^2)) / 6 + \\ \text{Herd/Sire 3: } & ((0^2) + (0^2) + (0^2) + (5^2)) / 5 + \\ \text{Herd/Sire 4: } & ((1^2) + (3^2) + (0^2) + (0^2)) / 4 + \\ \text{Herd/Sire 5: } & ((0^2) + (0^2) + (5^2) + (0^2)) / 5 + \\ \text{Herd/Sire 6: } & ((0^2) + (0^2) + (4^2) + (2^2)) / 6 + \\ \text{Herd/Sire 7: } & ((0^2) + (3^2) + (0^2) + (0^2)) / 3 + \\ \text{Herd/Sire 8: } & ((0^2) + (4^2) + (2^2) + (0^2)) / 6 + \\ \text{Herd/Sire 9: } & ((5^2) + (3^2) + (0^2) + (0^2)) / 8 + \\ \text{Herd/Sire 10: } & ((0^2) + (0^2) + (3^2) + (6^2)) / 9 \\ & = 37.35\end{aligned}$$

11. The $\sigma^2(\text{Herd})$ term on HS the sample size is calculated by the squared sum of:
cows that provided data across Herds for each Herd/Sire combination, divided by the total number of cows. This results in the N size.

$$= (57^2) / 57 = 57$$

12. The $\sigma^2(\text{Sire})$ term on HS the sample size is calculated by the squared sum of:
cows that provided data across Sires for each Herd/Sire combination, divided by the total number of cows. This results in the N size.

$$= (57^2) / 57 = 57$$

13. The $\sigma^2(\text{Year})$ term on Mean the sample size is calculated by the squared sum of:
cows in each year divided by the total number of cows.

$$= ((7^2) + (21^2) + (16^2) + (13^2))/57 = 16.05$$

14. The $\sigma^2(\text{Herd})$ term on Mean the sample size is calculated by the squared sum of:
#cows in each herd divided by the total number of cows.

$$= ((16^2) + (15^2) + (9^2) + (17^2))/57 = 14.93$$

15. The $\sigma^2(\text{Sire})$ term on Mean the sample size is calculated by the squared sum of:
#cows in each sire divided by the total number of cows.

$$= ((20^2) + (17^2) + (20^2))/57 = 19.11$$

16. The $\sigma^2(\text{HS})$ term on Mean the sample size is calculated by the squared sum of:
#cows in each Herd/Sire combination divided by the total number of cows.

$$= ((5^2) + (6^2) + (5^2) + (4^2) + (5^2) + (6^2) + (3^2) + (6^2) + (8^2) + (9^2))/57 = 6.19$$

All these calculated values can now be put into Table 2 and are presented in Table 3.

Table 3. Completed Hendersons T-values (uncorrected Sums of Squares) and coefficient sample size bases (contains the same data found in Table 6 p. 233 in Henderson 1953)

| Facet | T-Values | μ^2 | $\sigma^2(\text{year})$ | $\sigma^2(\text{herd})$ | $\sigma^2(\text{sire})$ | $\sigma^2(\text{HS})$ | $\sigma^2(\text{error})$ |
|-----------------|------------|---------|-------------------------|-------------------------|-------------------------|-----------------------|--------------------------|
| A (year) | 10,776,451 | 57 | 57 | 19.51 | 39.22 | 15.10 | 4 |
| H (herd) | 10,893,666 | 57 | 21.49 | 57 | 24.04 | 24.04 | 4 |
| S (sire) | 10,776,278 | 57 | 30.33 | 18.51 | 57 | 18.51 | 3 |
| HS (herdX sire) | 10,970,369 | 57 | 37.35 | 57 | 57 | 57 | 10 |
| Mean | 10,685,141 | 57 | 16.05 | 14.93 | 19.11 | 6.19 | 1 |
| Total | 11,124,007 | 57 | 57 | 57 | 57 | 57 | 57 |

Although at this point Henderson creates a “Table 7” from which equations he says can be solved, this step is not necessary. What is necessary is to “solve” for the equations that are set up in Table 3. Specifically, what values need to be multiplied with each estimated coefficient down the entire column for all 6 equations that will solve the following equations? These will be the variance components: A, B, C, D, E, and F.

$$10,776,451 = A(57) + B(57) + C(19.51) + D(39.22) + E(15.10) + F(4)$$

$$10,893,666 = A(57) + B(21.49) + C(57) + D(24.04) + E(24.04) + F(4)$$

$$10,776,278 = A(57) + B(30.33) + C(18.51) + D(57) + E(18.51) + F(3)$$

$$10,970,369 = A(57) + B(37.35) + C(57) + D(57) + E(57) + F(10)$$

$$10,685,141 = A(57) + B(16.05) + C(14.93) + D(19.11) + E(6.19) + F(1)$$

$$11,124,007 = A(57) + B(57) + C(57) + D(57) + E(57) + F(57)$$

Using the data in Table 3, solve these equations simultaneously via matrix procedures to solve a system of equations (solving by linear least squares method). Alternatively, they can be solved by regressing the 6 T-values on the 6 variance estimates in a direct regression (all predictors entered simultaneously), and requesting to include an intercept in the model). This produces the following output (from SPSS – although other programs will provide the same results). Note that there is no “Mean” variance component calculated, as μ^2 is a constant and has no variance so is eliminated from the analysis. Since there are 6 predictors and 6 outcomes the data will “fit perfectly” (in fact, Excel WILL NOT run the regression because there are equal numbers of predictors and cell entries in the outcome). However, this is not relevant, as we are interested solely in the unstandardized B-values associated with each effect, which are the variance components. We now have the variance components (see Table 4) for Henderson’s data.

| Model | | Unstandardized Coefficients | | Standardized Coefficients | | |
|-------|----------------|-----------------------------|------------|---------------------------|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 10572974.498 | .000 | . | . | . |
| | Year_tutorial | 763.154 | .000 | .084 | . | . |
| | Herd_tutorial | 4531.304 | .000 | .615 | . | . |
| | Sire_tutorial | 1587.278 | .000 | .174 | . | . |
| | HS_tutorial | -164.329 | .000 | -.023 | . | . |
| | Error_tutorial | 2949.830 | .000 | .402 | . | . |

Table 4. Variance Components from Henderson 1953 data

| Facet | Variance Component | Proportion of Variance |
|----------|--------------------|------------------------|
| A (year) | 763 | .08 |
| H (herd) | 4531 | .46 |
| S (sire) | 1587 | .16 |
| HS | -164 | .30 |
| Error | 2950 | |
| Total | | |

We can see from the results that most of the variance (46%) in the butterfat amount, is due to the “Herd” variable and next by the Herd X Sire interaction (30%). Not a lot of the variance is due to the Sire (16%) and almost none is due to Year (8%). As noted earlier, Henderson stops at this point.

Before leaving Henderson, there are a couple notes about his protocol.

1. It is obvious now why this approach allows for missing data and unbalanced designs in the estimate of variance components. While tedious, the calculation of the T-values and sample sizes for effects, do not require balanced, non-missing data sets that meet all assumptions of normality. This flexibility is a marvelous feature insofar as naturally occurring data sets often fall short of these restrictive assumptions.
2. ALL effects are assumed to be random. If an effect (such as “year”) IS actually fixed, then the other estimates are biased. This assumption continues into Brennan’s (2001a) work, and thus the into the urGENOVA and G-String_V programs.
3. The variance components are not used to create generalizability coefficients in Henderson’s article. They can be, but the generalizability literature had not begun to be published until after this article. Brennan (2001a) took up the calculations to extend generalizability’s reach into data sets that can be nested, unbalanced and/or have missing data.

Brennan’s (2001a) approach (he calls it analogous-ANOVA) is based on Henderson’s calculations of T-values. With unbalanced or missing data, “The primary theoretical problem presented by most unbalanced designs is that there are many possible estimators of random effects variance components and no unambiguously clear basis for determining which estimators are best. Among the practical problems with unbalanced designs are that some estimation methods require distributional form assumptions, which are often difficult to justify in generalizability analyses.” (Manual for urGENOVA Brennan, 2001b p. 1). Thus, he advises using the variance components generated from urGENOVA for generalizability estimates that are strictly “descriptive” of the specific data set and not inferential. The urGENOVA program does not allow for any computations for D-studies, nor are standard errors and confidence intervals of the variance components appropriate except for balanced designs under the assumptions of normality. A shortcoming of the urGENOVA program is that IT DOES NOT produce generalizability or dependability coefficients. These have to be calculated by hand by the end user.

The G-String_V program is more restrictive in input than urGENOVA. It does NOT allow for missing data. This simplifies the calculations (as we will see in a moment). The authors of this program do not adhere to Brennan’s (2001b) admonition of D-study, standard errors and confidence intervals of the variance components. In fact, it is relatively straightforward in the case of unbalanced, but no missing data, to estimate D-study generalizability coefficients using harmonic mean estimates of the

unbalanced effects. However, these should be taken as estimates only, given the issues raised by Brennan. The G-String_V program DOES calculate generalizability and dependability coefficients for the facet of differentiation specified by the user.

Narayanan et al. (2010) and Brennan (2001a)

With this as a backdrop, the next step in this tutorial is to use a data set from Narayanan et al. (2010) (Table 5) to calculate the variance components using Henderson's Method 1. It will be noted at this point that Narayanan uses a completely different approach to calculating generalizability in the paper. It is based on the variance, not the variance components, which is the hallmark of traditional approaches to calculating Generalizability. These data are unbalanced, but no data points are missing. There are 3 different doctors (d); rated by 16 patients (p) - note that patients are nested in doctors, and this facet is unbalanced; 8 patients rate Doctor A, 5 rate Doctor B and 3 rate Doctor C); all use 5 items (i) to rate the doctor. The model then is: $i \times X(p:d)$. The dependent variable is "rating on a scale of 1-5" for each item.

Table 5. Narayanan et al. (2010) data set ($i \times p:d$)

| Doctor | Patient | item1 | item2 | item3 | item4 | item5 |
|--------|---------|-------|-------|-------|-------|-------|
| A | 1 | 4 | 4 | 4 | 4 | 4 |
| A | 2 | 4 | 4 | 3 | 3 | 4 |
| A | 3 | 3 | 4 | 4 | 3 | 4 |
| A | 4 | 3 | 3 | 3 | 3 | 3 |
| A | 5 | 3 | 3 | 4 | 4 | 3 |
| A | 6 | 4 | 4 | 3 | 3 | 3 |
| A | 7 | 3 | 3 | 3 | 3 | 3 |
| A | 8 | 4 | 4 | 3 | 3 | 3 |
| B | 9 | 4 | 4 | 4 | 4 | 3 |
| B | 10 | 4 | 4 | 4 | 4 | 4 |
| B | 11 | 4 | 4 | 4 | 4 | 4 |
| B | 12 | 4 | 3 | 3 | 3 | 3 |
| B | 13 | 4 | 4 | 3 | 3 | 3 |
| C | 14 | 4 | 4 | 3 | 3 | 3 |
| C | 15 | 3 | 3 | 3 | 3 | 3 |
| C | 16 | 4 | 4 | 4 | 4 | 3 |

Step 1: Calculating the T-values (uncorrected sums of squares) for each facet

We are going to need a number of sums and their squares for the next analyses (Table 6):

Table 6. Narayanan et al. (2010) data set with added rows and columns

| Doctor | Patient | item1 | item2 | item3 | item4 | item5 | Sum across Patient Ratings | Squared sum across Patient Ratings | Squared sum across patient ratings/#ratings each completed |
|--------|--|----------|----------|----------|--------|----------|----------------------------|------------------------------------|--|
| A | 1 | 4 | 4 | 4 | 4 | 4 | 20 | 400 | 80 |
| A | 2 | 4 | 4 | 3 | 3 | 4 | 18 | 324 | 64.8 |
| A | 3 | 3 | 4 | 4 | 3 | 4 | 18 | 324 | 64.8 |
| A | 4 | 3 | 3 | 3 | 3 | 3 | 15 | 225 | 45 |
| A | 5 | 3 | 3 | 4 | 4 | 3 | 17 | 289 | 57.8 |
| A | 6 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| A | 7 | 3 | 3 | 3 | 3 | 3 | 15 | 225 | 45 |
| A | 8 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| B | 9 | 4 | 4 | 4 | 4 | 3 | 19 | 361 | 72.2 |
| B | 10 | 4 | 4 | 4 | 4 | 4 | 20 | 400 | 80 |
| B | 11 | 4 | 4 | 4 | 4 | 4 | 20 | 400 | 80 |
| B | 12 | 4 | 3 | 3 | 3 | 3 | 16 | 256 | 51.2 |
| B | 13 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| C | 14 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| C | 15 | 3 | 3 | 3 | 3 | 3 | 15 | 225 | 45 |
| C | 16 | 4 | 4 | 4 | 4 | 3 | 19 | 361 | 72.2 |
| | Item Rating Sums | 59 | 59 | 55 | 54 | 53 | | | |
| | Item Rating sums squared | 3481 | 3481 | 3025 | 2916 | 2809 | | | |
| | Item sums squared/#ratings For each item | 217.5625 | 217.5625 | 189.0625 | 182.25 | 175.5625 | | | |

1. Doctor (d)

Sum all ratings across each item for each Patient; Sum the individual patient ratings for each doctor; square these sums; divide by the number of rating counts within each doctor; add the quotients.

$$\text{Doctor A} = (20+18+18+15+17+17+15+17) = 137; 137^2 = 18769; 18769/40 = 469.225$$

$$\text{Doctor B} = (19+20+20+16+17) = 92; 92^2 = 8464; 8464/25 = 338.560$$

$$\text{Doctor C} = (17+15+19) = 51 = 51^2 = 2601; 2601/15 = 173.400$$

$$\text{Sum across all 3 quotients} = 469.225 + 338.560 + 173.400 = 981.185$$

2. Patient:Doctor (p:d)

Sum all ratings across each item for each Patient; Square these sums; divide by the number of rating counts within each patient; sum across these quotients.

$$p:d 1 (20^2) = 400; 400/5 = 80$$

$$p:d 2 (18^2) = 324; 324/5 = 64.8$$

.....

$$p:d 15 (15^2) = 225; 225/5 = 45$$

$$p:d 16 (19^2) = 362; 361/5 = 72.2$$

$$\text{Sum across all 16 quotients} = 989.200$$

3. Item (i)

Sum all ratings down each item; Square these sums; divide by the number of rating counts within each item; sum across these quotients.

$$\text{Item 1} (59^2) = 3481; 3481/16 = 217.5625$$

$$\text{Item 2} (59^2) = 3481; 3481/16 = 217.5625$$

$$\text{Item 3} (55^2) = 3025; 3025/16 = 189.0625$$

$$\text{Item 4} (54^2) = 2916; 2916/16 = 182.250$$

$$\text{Item 5} (53^2) = 2809; 2809/16 = 175.5625$$

$$\text{Sum across all 5 quotients} (217.5625+....+175.5625) = 980.000$$

4. Doctor X Item (di)

Sum down each set of items for each doctor (there will be 15 combinations); square these sums; divide each sum by the number of ratings that go into each of the 15 combinations; sum the quotients.

$$di1 = (4+4+3+3+3+4+3+4) = 28; 28^2 = 784; 784/8 = 98.000$$

$$di2 = (4+4+4+3+3+4+3+4) = 29; 29^2 = 841; 841/8 = 105.125$$

.....

$$di6 = (4+4+4+4+4) = 20; 20^2 = 400; 400/5 = 80.000$$

$$di7 = (4+4+4+3+4) = 19; 19^2 = 361; 361/5 = 72.200$$

.....

$$di14 = (3+3+4) = 10; 10^2 = 100; 100/3 = 33.333$$

$$di15 = (3+3+3) = 9; 9^2 = 81; 81/3 = 27.000$$

$$\text{Sum across all quotients} = 983.808$$

5. PatientXItem:Doctor (pi:d) = TOTAL*

Note that this is the total uncorrected sums of squares. Each individual rating is first squared. Then these are summed across all ratings.

Going across each row....

$$4^2 + 4^2 + 4^2 + 4^2 + 4^2 +$$

$$4^2 + 4^2 + 4^2 + 4^2 + 3^2 +$$

....continuing for all rows

$$3^2 + 3^2 + 3^2 + 3^2 + 3^2 +$$

$$4^2 + 4^2 + 4^2 + 4^2 + 3^2$$

$$\text{Summing across all values} = 1000.000$$

6. Mean

Sum FIRST across all values; square this sum; divide by the total number of ratings that went into the sum.

4 + 4+4+ 4+4 +

4 +4 + 4 +4 + 3 +

....continuing for all rows

3 + 3 +3 + 3 +3 +

4 +4 + 4 +4 + 3

Summing across all values = 280; $280^2 = 78400$; $78400/80 = 980.000$

We can now put our Facet T-values into Table 7. below.

Table 7. Facets and T-values for the Narayanan et al. (2010) data set

| Facet | T-Value |
|---|----------|
| D (doctor) | 981.185 |
| P:D (patient nested in doctor) | 989.200 |
| I (item) | 982.000 |
| DI (doctor X item) | 983.808 |
| PI:D (patient X item nested in doctor) (Equivalent to Henderson's "total") | 1000.000 |
| Mean | 980.000 |

Step 2: Calculating variance coefficients for each facet

As we know from the Henderson example, the next part is tedious. We have to get the sample sizes for the variance coefficients. Like Henderson, we will first put in the "easy" ones.

1. For the μ^2 term, ALL coefficients are based on the total number of cases involved in the study (denoted N; 80 for this data set).
2. For the PI:D (Total which includes the highest level term plus error) facet, all coefficients are also based on the total number of cases involved in the study (denoted N).
3. For each facet (except the Mean), the coefficient for that variance is also the total number of cases involved in the study (denoted N). For example, the coefficient for $\sigma^2(d)$ for effect D is N.
4. For the $\sigma^2(\pi:d)$ term column, the sample sizes are equal to the number of levels of that facet (D = 3, P:D = 16, I = 5, DI = 15 (there are 15 different DI combinations), for the Mean is always = 1, and for the highest order effect (PI:D – also known as the total by Henderson) is always = N.

Brennan (2001a, equations 7.2-7.6 p. 219) discusses how each of the cell entries are generated. These equations are difficult to "unpack" so I will be using Henderson's "brute force" approach that allows for every count to be taken into consideration when calculating the terms.

When the data set is not missing any datapoints, Brennan uses a notation system for the coefficients in the other cells that need to be calculated. These are mostly simple, such as $n(p)$ is the number of levels of the facet. In this design the $n(d) = 3$; $n(p) = 16$; $n(i) = 5$. However, $n(r)$ is a unique notation. It is based on the number of "counts(individual ratings)" within each DI combination (recall there are 15 of them); square these values; sum them; divide by the number of counts in the entire data set.

$$((8^2) + (8^2) + (8^2) + (8^2) + (8^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (3^2) + (3^2) + (3^2) + (3^2) + (3^2))/80 = 6.125$$

The results from Brennan's notation are shown in Table 8 for this design. However, as noted, the entries ONLY work in a balanced design. When there are missing data, we need to go back to the Henderson approach.

Table 8. Facets and Variance Coefficients for Narayanan et al. (2010) data set with Brennan (2001a) Notation

| Facet | $\sigma^2(d)$ | $\sigma^2(p:d)$ | $\sigma^2(i)$ | $\sigma^2(di)$ | $\sigma^2(pi:d)$ | μ^2 |
|----------------------------|---------------|-----------------|---------------|----------------|------------------|---------|
| D | N | $n(i)*n(d)$ | $n(p)$ | $n(p)$ | $n(d)$ | N |
| P:D | N | N | $n(p)$ | $n(p)$ | $n(p)$ | N |
| I | $n(r)*n(i)$ | $n(i)$ | N | $n(r)*n(i)$ | $n(i)$ | N |
| DI | N | $n(i)*n(d)$ | N | N | $n(i)*n(d)$ | N |
| PI:D (total for Henderson) | N | N | N | N | N | N |
| Mean | $n(r)*n(i)$ | $n(i)$ | $n(p)$ | $n(r)$ | 1 | N |

Henderson's approach to completing the "non-simple" cells in the coefficient matrix follows. It allows for missing data in the calculations. There are 16 unique values (highlighted) to calculate.

Table 9. Facets and Variance Coefficients for Narayanan et al. (2010) data set with needed values 1-16 noted

| Facet | $\sigma^2(d)$ | $\sigma^2(p:d)$ | $\sigma^2(i)$ | $\sigma^2(di)$ | $\sigma^2(pi:d)$ | μ^2 |
|----------------------------|---------------|-----------------|---------------|----------------|------------------|---------|
| D | N | 1 | 2 | 3 | $n(d)$ | N |
| P:D | 4 | N | 5 | 6 | $n(p)$ | N |
| I | 7 | 8 | N | 9 | $n(i)$ | N |
| DI | 10 | 11 | 12 | N | $n(i)*n(d)$ | N |
| PI:D (total for Henderson) | N | N | N | N | N | N |
| Mean | 13 | 14 | 15 | 16 | 1 | N |

1. The $\sigma^2(p:d)$ term on D (data are collapsed across Items). Sample size starts by calculating the squared sum of:
counts of each p:d term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2))/40 = 200/40 = 5 +$$

$$D2: ((5^2) + (5^2) + (5^2) + (5^2) + (5^2))/25 = 125/25 = 5 +$$

$$D3: ((5^2) + (5^2) + (5^2))/15 = 75/15 = 5$$

15

2. The $\sigma^2(i)$ term on D (data are collapsed across P:D). Sample size starts by calculating the squared sum of:
counts of each i term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((8^2) + (8^2) + (8^2) + (8^2) + (8^2))/40 = 320/40 = 8 +$$

$$D2: ((5^2) + (5^2) + (5^2) + (5^2) + (5^2))/25 = 125/25 = 5 +$$

$$D3: ((3^2) + (3^2) + (3^2) + (3^2) + (3^2))/15 = 45/15 = 3$$

16

3. The $\sigma^2(di)$ term on D (data are collapsed across P:D). Sample size starts by calculating the squared sum of:
counts of each di term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((8^2) + (8^2) + (8^2) + (8^2) + (8^2))/40 = 320/40 = 8 +$$

$$D2: ((5^2) + (5^2) + (5^2) + (5^2) + (5^2))/25 = 125/25 = 5 +$$

$$D3: ((3^2) + (3^2) + (3^2) + (3^2) + (3^2))/15 = 45/15 = 3$$

16

4. The $\sigma^2(d)$ term on P:D (data are collapsed across Items). Sample size starts by calculating the squared sum of:
counts of each d term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

$$PD\ 1: (5^2)/5 = 5 +$$

$$PD\ 2: (5^2)/5 = 5 +$$

$$PD\ 3: (5^2)/5 = 5 +$$

$$PD\ 4: (5^2)/5 = 5 +$$

$$PD\ 5: (5^2)/5 = 5 +$$

PD 6: $(5^2)/5 = 5 +$
 PD 7: $(5^2)/5 = 5 +$
 PD 8: $(5^2)/5 = 5 +$
 PD 9: $(5^2)/5 = 5 +$
 PD 10: $(5^2)/5 = 5 +$
 PD 11: $(5^2)/5 = 5 +$
 PD 12: $(5^2)/5 = 5 +$
 PD 13: $(5^2)/5 = 5 +$
 PD 14: $(5^2)/5 = 5 +$
 PD 15: $(5^2)/5 = 5 +$
 PD 16: $(5^2)/5 = 5$
 =80

5. The $\sigma^2(i)$ term on P:D. Sample size starts by calculating the squared sum of:
 # counts of each i term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

PD 1: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 2: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 3: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 4: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 5: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 6: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 7: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 8: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 9: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 10: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 11: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 12: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 13: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 14: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 15: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 16: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1$
 =16

6. The $\sigma^2(di)$ term on P:D. Sample size starts by calculating the squared sum of:
 # counts of each di term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

PD 1: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 2: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 3: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 4: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 5: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 6: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 7: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 8: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 9: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 10: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 11: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 12: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 13: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 14: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 15: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$

$$\text{PD 16: } ((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 \\ =16$$

7. The $\sigma^2(d)$ term on I. Sample size starts by calculating the squared sum of:
counts of each d term within each I, divided by the total number of counts for that I. Then sum the quotients.

$$\begin{aligned} \text{I 1: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 2: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 3: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 4: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 5: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 \\ & =30.625 \end{aligned}$$

8. The $\sigma^2(p:d)$ term on I. Sample size starts by calculating the squared sum of:
counts of each p:d term within each I, divided by the total number of counts for that I. Then sum the quotients.

$$\begin{aligned} \text{I 1: } & ((1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2))/16 = 16/16 = 1+ \\ \text{I 2: } & ((1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2))/16 = 16/16 = 1+ \\ \text{I 3: } & ((1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2))/16 = 16/16 = 1+ \\ \text{I 4: } & ((1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2))/16 = 16/16 = 1+ \\ \text{I 5: } & ((1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2)+(1^2))/16 = 16/16 = 1 \\ & =5 \end{aligned}$$

9. The $\sigma^2(di)$ term on I. Sample size starts by calculating the squared sum of:
counts of each di term within each I, divided by the total number of counts for that I. Then sum the quotients.

$$\begin{aligned} \text{I 1: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 2: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 3: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 4: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 + \\ \text{I 5: } & ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 \\ & =30.625 \end{aligned}$$

10. The $\sigma^2(d)$ term on DI. Sample size starts by calculating the squared sum of:
counts of each d term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

$$\begin{aligned} \text{DI 1: } & (8^2)/8 = 8 + \\ \text{DI 2: } & (8^2)/8 = 8 + \\ \text{DI 3: } & (8^2)/8 = 8 + \\ \text{DI 4: } & (8^2)/8 = 8 + \\ \text{DI 5: } & (8^2)/8 = 8 + \\ \text{DI 6: } & (5^2)/5 = 5 + \\ \text{DI 7: } & (5^2)/5 = 5 + \\ \text{DI 8: } & (5^2)/5 = 5 + \\ \text{DI 9: } & (5^2)/5 = 5 + \\ \text{DI 10: } & (5^2)/5 = 5 + \\ \text{DI 11: } & (3^2)/3 = 3 + \end{aligned}$$

DI 12: $(3^2)/3 = 3 +$
 DI 13: $(3^2)/3 = 3 +$
 DI 14: $(3^2)/3 = 3 +$
 DI 15: $(3^2)/3 = 3$
 =80

11. The $\sigma^2(p:d)$ term on DI. Sample size starts by calculating the squared sum of:

counts of each p:d term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

DI 1: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$
 DI 2: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$
 DI 3: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$
 DI 4: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$
 DI 5: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$
 DI 6: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$
 DI 7: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$
 DI 8: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$
 DI 9: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$
 DI 10: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$
 DI 11: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 12: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 13: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 14: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 15: $((1^2) + (1^2) + (1^2))/3 = 1$
 =15

12. The $\sigma^2(i)$ term on DI. Sample size starts by calculating the squared sum of:

counts of each i term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

DI 1: $(8^2)/8 = 8 +$
 DI 2: $(8^2)/8 = 8 +$
 DI 3: $(8^2)/8 = 8 +$
 DI 4: $(8^2)/8 = 8 +$
 DI 5: $(8^2)/8 = 8 +$
 DI 6: $(5^2)/5 = 5 +$
 DI 7: $(5^2)/5 = 5 +$
 DI 8: $(5^2)/5 = 5 +$
 DI 9: $(5^2)/5 = 5 +$
 DI 10: $(5^2)/5 = 5 +$
 DI 11: $(3^2)/3 = 3 +$
 DI 12: $(3^2)/3 = 3 +$
 DI 13: $(3^2)/3 = 3 +$
 DI 14: $(3^2)/3 = 3 +$
 DI 15: $(3^2)/3 = 3$
 =80

13. The $\sigma^2(d)$ term on mean. Sample size starts by calculating the squared sum of:

counts of each d term within each mean, divided by the total number of counts for the mean.

Mean: $((40^2) + (25^2) + (15^2))/80 = (1600 + 625 + 225)/80 = 2450/80 = 30.625$

14. The $\sigma^2(p:d)$ term on mean. Sample size starts by calculating the squared sum of:
counts of each p:d term within each mean, divided by the total number of counts for the mean.

$$\text{Mean: } ((5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2))/80 = (25*16)/80 = 400/80 = 5$$

15. The $\sigma^2(i)$ term on mean. Sample size starts by calculating the squared sum of:
counts of each i term within each mean, divided by the total number of counts for the mean

$$\text{Mean: } ((16^2) + (16^2) + (16^2) + (16^2) + (16^2))/80 = 256*5/80 = 1280/80 = 16$$

16. The $\sigma^2(di)$ term on mean. Sample size starts by calculating the squared sum of:
counts of each di term within each mean, divided by the total number of counts for the mean.

$$\begin{aligned} \text{Mean: } & ((8^2) + (5^2)+(3^2)+(8^2)+(5^2)+(3^2)+(8^2)+(5^2)+(3^2)+(8^2)+(5^2)+(3^2)+(8^2)+(5^2)+(3^2))/80 \\ & = (64*5 + 25*5 + 9*5)/80 \\ & = (320 + 125 + 45)/80 \\ & = 490/80 = 6.125 \end{aligned}$$

We insert the calculated values into their respective cells (Table 10) and bring the T-values down to complete the table.

Table 10. Facet and Variance Coefficients for Narayanan et al. (2010) data set

| Facet | $\sigma^2(d)$ | $\sigma^2(p:d)$ | $\sigma^2(i)$ | $\sigma^2(di)$ | $\sigma^2(pi:d)$ | μ^2 | T-Value |
|----------------------------|---------------|-----------------|---------------|----------------|------------------|---------|----------|
| D | 80 | 15 | 16 | 16 | 3 | 80 | 981.185 |
| P:D | 80 | 80 | 16 | 16 | 16 | 80 | 989.200 |
| I | 30.625 | 5 | 80 | 30.625 | 5 | 80 | 982.000 |
| DI | 80 | 15 | 80 | 80 | 15 | 80 | 983.808 |
| PI:D (total for Henderson) | 80 | 80 | 80 | 80 | 80 | 80 | 1000.000 |
| Mean | 30.625 | 5 | 16 | 6.125 | 1 | 80 | 980.000 |

Using the data in Table 10, (via matrix procedures or regressing the T-values on the coefficients) the variance components for the facets are calculated. Below is the output from SPSS. Because μ^2 is a constant and has no variance, it is eliminated from the regression analysis. However, the Constant (978.972) when divided by the sample size (80) is 12.237, which is the mean variance component if matrix procedures had been used.

| Model | | Unstandardized Coefficients | | Standardized Coefficients |
|-------|------------|-----------------------------|------------|---------------------------|
| | | B | Std. Error | Beta |
| 1 | (Constant) | 978.972 | .000 | |
| | D | .002 | .000 | .008 |
| | P_D | .092 | .000 | .442 |
| | I | .028 | .000 | .128 |
| | DI | -.016 | .000 | -.071 |
| | PI_D_E | .157 | .000 | .625 |

Examination of the variance components of the facets and their relative contribution to variance away from the mean (Table 11), we can see that most of the variance (56%) in this data set is due to the PI:D facet. This is problematic from an interpretive standpoint, indicating it is not possible to really understand the variance in the data set. 33% of the variance is due to the P:D facet – within patient ratings of the doctors. In examining the raw data, we see that the doctors are all rated similarly by all patients (very little variance in the data set with which to work).

Table 11. Variance Components from Narayanan et al. (2010) data

| Facet | Variance Component | Proportion of Variance |
|-------|--------------------|------------------------|
| D | .002 | .007 |
| P:D | .092 | .330 |
| I | .028 | .100 |
| DI | -.016** | 0 |
| PI:D | .157 | .563 |

Calculating Generalizability and Dependability (G and D) Coefficients

Using the Variance Component (VC) values, Generalizability (G) and Dependability (D) coefficients can be calculated. It is most likely that we would want to know “Are the patients reliable in making their ratings of the doctors?” That is, “patient” would be the face of differentiation. We might also want to know how reliable the items are so we can also use item as a facet of differentiation. We can also specify doctor as a facet of differentiation, but looking at the data, there is so little variance on this facet that the results will not likely be reliable.

To calculate G and D, we need the variance components for each facet **as well as the denominators** for each of the facets that will serve in the error term (the facet(s) of generalization) when calculating the G and D for each facet of differentiation. Note that these denominators will differ depending on which facet is the facet of differentiation, and which facet/facets are being generalized across.

The calculations for the denominators are extensions on Brennan’s (2001a, p. 232) equation 7.28. The calculation modifies equation 7.28 for each unique value within the facet of differentiation and returns the harmonic mean* (see Appendix A for a description of the use of harmonic mean) of those values to give the appropriate level coefficient. The general formula for determining levels used for G Theory Calculations can be determined as follows:

Let:

- G = the set of grouping combinations for the facet of differentiation. For example, if the facet of differentiation is “doctors”, the set of grouping combinations would simply be unique values of doctors (A, B, C), whereas if the facet of differentiation is “patients:doctors” the set of grouping combinations would be each unique combinations of patients and doctors [(1, A), (2, A), ..., (16, C)].
- V = the set of grouping combinations for the variance over which generalization is occurring. For example, if we are generalizing the facet of p:d across items, we will want to determine the levels for variance(items), given the facet of differentiation p:d.
- C(g) = count of occurrences of the dependent variable for the grouping combination, $g \in G$

For each g, the level is determined by the sum of counts squared over the sum of square counts for $v \in V$:

$$L_g = \frac{(\sum_{v=1}^V c(g,v))^2}{\sum_{v=1}^V c(g,v)^2}$$

When the data set is balanced and complete (i.e. no missing terms) this expression simplifies to $|V|$, which is equal to the counts of unique combinations of facets in the variance in question for a given facet of differentiation.

It should be straightforward to the reader to see why this is the case. Let’s imagine briefly that for this design [items x (patients:doctors)] we had a balanced nesting of 5 items, 3 doctors, and 5 patients per doctor (instead of 8, 5, and 3 patients respectively). If the facet of differentiation was ‘doctors’ and we were trying to determine the appropriate level for ‘patients:doctors’ the calculation would be as follows:

$$L_g = \frac{(\sum_{v=1}^5 5)^2}{\sum_{v=1}^5 5^2} = \frac{(5*5)^2}{5*5^2} = 5$$

Here each $C(g,v) = 5$ because there are 5 items for each combination of patient:doctor, and since there are 5 patients in the balanced example $v \in [1, 5]$. Thus, in the balanced case where there are 5 patients per doctor, it makes sense that we must correct the variance(patient:doctor) by a factor of 5 when we are considering doctors as the facet of differentiation!

Notice that thus far we've actually only calculated the level, L , for a unique facet of differentiation, g . When the data set is balanced, or we are dealing with an unnested variable (i.e. "items" is unnested in our current [items x (patients:doctors)] design) in an unbalanced design, each unique g will have the same level, L . For example, consider the unbalanced dataset in question. If we are considering 'items' as the facet of differentiation, and calculating the levels for 'p:d', it should be clear that each of the five items, g , has 16 unique combinations. However, this is not the case for facets involved with unbalanced or missing data. We must extend our definition of L to account for all $g \in G$ by taking the harmonic mean of each L_g :

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{|G|}{\sum_{g=1}^G \frac{\sum_{v=1}^V C(g,v)^2}{\left(\sum_{v=1}^V C(g,v)\right)^2}}$$

The harmonic mean is necessary as we are accounting for the *ratio* between the counts of the variance in question and the chosen facet of differentiation. When data are balanced and complete this expression is unnecessary as $g_i = g_k$ for all $g \in G$ and thus $L = L_g$. However, this is not the case for unbalanced and missing data, so we must correct the variance by the harmonic mean of all L_g , L .

Table 12 shows the denominators to be used in our unbalanced data set with the complete worked examples following.

Table 12. Denominators for each effect when the facet of differentiation changes when calculating G and D

| Variance Term | P:D Facet of Differentiation Denominator for G and D | I Facet of Differentiation Denominator for G and D | D Facet of Differentiation Denominator for G and D |
|----------------------------|---|---|---|
| $\sigma^2(d)$ | 1 | 2.612 | 1 |
| $\sigma^2(p:d)$ | 1 | 16 | 4.557 |
| $\sigma^2(i)$ | 5 | 1 | 5 |
| $\sigma^2(i \times d)$ | 5 | 2.612 | 5 |
| $\sigma^2(i \times (p:d))$ | 5 | 16 | 22.785 |

1. Level of the $\sigma^2(d)$ term on facet of differentiation P:D. Since d is included in P:D, there is a single unique D for each P:D. $L = 1$

2. Level of the $\sigma^2(p:d)$ term on facet of differentiation P:D. Again, $p:d$ is included in P:D, so there is a single unique P:D for each P:D. $L = 1$

3. Level of the $\sigma^2(i)$ term on facet of differentiation P:D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(p:d, i) = 1, v = 5$: 1 count of item for each unique combination of $p:d$ and i with $i \in [1, 5]$ for each $p:d$.

$$L = L_g = \frac{\left(\sum_{v=1}^5 1\right)^2}{\sum_{v=1}^5 1^2} = 5$$

4. Level of the $\sigma^2(i \times d)$ term on facet of differentiation P:D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(p:d, i \times d) = 1, v = 5$: Again there is 1 count of item x doctor for each unique combination of $p:d$ and $i \times d$ (notice that d occurs in both) with $i \times d \in [1, 5]$ for each $p:d$.

$$L = L_g = \frac{(\sum_{v=1}^5 1)^2}{\sum_{v=1}^5 1^2} = 5$$

5. Level of the $\sigma^2(i \times (p:d))$ term on facet of differentiation P:D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(p:d, i \times (p:d)) = 1$, $v = 5$: Again there is 1 count of $i \times (p:d)$ for each unique combination of $p:d$ and $i \times (p:d)$ (notice both d and p occurs in both) with $i \times (p:d) \in [1, 5]$ for each $p:d$.

$$L = L_g = \frac{(\sum_{v=1}^5 1)^2}{\sum_{v=1}^5 1^2} = 5$$

6. Level of the $\sigma^2(d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, d) = |p_v|$, for $p = [8, 5, 3]$ with $v \in [1, 3]$; $g \in [1, 5]$. Since i is not involved with the unbalanced nesting, all levels of i , $i \in [1, 5]$, will be the same.

$$L = L_{1-5} = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+3)^2}{(8^2+5^2+3^2)} = \frac{256}{98} = 2.612$$

7. Level of the $\sigma^2(p:d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, p:d) = 1$, $v \in [1, 16]$. For each unique i , there will be 16 unique counts of $p:d$.

$$L = L_{1-5} = \frac{(\sum_{v=1}^{16} 1)^2}{\sum_{v=1}^{16} 1^2} = 16$$

8. Level of the $\sigma^2(i)$ term on facet of differentiation I. Since i is included in I, there is a single unique i for each I. $L = 1$

9. Level of the $\sigma^2(i \times d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, i \times d) = |p_v|$, for $p = [8, 5, 3]$ with $v \in [1, 3]$; $g \in [1, 5]$. Since i is not involved with the unbalanced nesting, all levels of i , $i \in [1, 5]$, will be the same.

$$L = L_{1-5} = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+3)^2}{(8^2+5^2+3^2)} = \frac{256}{98} = 2.612$$

10. Level of the $\sigma^2(i \times (p:d))$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, I \times (p:d)) = 1$, $v \in [1, 16]$. For each unique i , there will be 16 unique counts of $i \times (p:d)$.

$$L = L_{1-5} = \frac{(\sum_{v=1}^{16} 1)^2}{\sum_{v=1}^{16} 1^2} = 16$$

11. Level of the $\sigma^2(d)$ term on facet of differentiation D. Since d is included in D, there is a single unique D for each D. $L = 1$

12. Level of the $\sigma^2(p:d)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, p:d) = 5$, for $[g=A, v \in [1, 8]; g=B, v \in [1, 5]; g=C, v \in [1, 3]]$. Thus we have different levels of patients:doctors for each unique doctor.

$$L_A = \frac{(\sum_{v=1}^8 5)^2}{\sum_{v=1}^8 5^2} = 8$$

$$L_B = \frac{(\sum_{v=1}^5 5)^2}{\sum_{v=1}^5 5^2} = 5$$

$$L_C = \frac{(\sum_{v=1}^3 5)^2}{\sum_{v=1}^3 5^2} = 3$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{8} + \frac{1}{5} + \frac{1}{3}} = 4.557$$

13. Level of the $\sigma^2(i)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i) = |g|$, for $g=[8, 5, 3]$ and $v = 5$. There are g counts of each item $[1, 5]$. For example, there are 8 counts of items $[1, 5]$ for doctor, d, A because there are 8 instances of 8 (by marginalizing over patients).

$$L_A = \frac{(\sum_{v=1}^5 8)^2}{\sum_{v=1}^5 8^2} = 5$$

$$L_B = \frac{(\sum_{v=1}^5 5)^2}{\sum_{v=1}^5 5^2} = 5$$

$$L_C = \frac{(\sum_{v=1}^5 3)^2}{\sum_{v=1}^5 3^2} = 5$$

$$L = L_C = L_B = L_A = 5$$

14. Level of the $\sigma^2(i \times d)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i \times d) = |p|$, for $p=[8, 5, 3]$ and $v = 5$. There are p counts of each item, $v, [1, 5]$. This is identical to the uncrossed items.

$$L_A = \frac{(\sum_{v=1}^5 8)^2}{\sum_{v=1}^5 8^2} = 5$$

$$L_B = \frac{(\sum_{v=1}^5 5)^2}{\sum_{v=1}^5 5^2} = 5$$

$$L_C = \frac{(\sum_{v=1}^5 3)^2}{\sum_{v=1}^5 3^2} = 5$$

$$L = L_C = L_B = L_A = 5$$

15. Level of the $\sigma^2(i \times (p:d))$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i \times (p:d)) = 1$, for $v=[p_g \times 5]$ with $p=[8, 5, 3]$. For each doctor, there is a single unique count of items crossed with the nested patient:doctor. Thus, to obtain the level for each unique facet of differentiation, g , we must sum over v , which is patients * items or $p \times 5$ (since patients is unbalanced under doctors and items is constant 5).

$$L_A = \frac{(\sum_{v=1}^{8 \times 5} 1)^2}{\sum_{v=1}^{8 \times 5} 1^2} = \frac{1600}{40} = 40$$

$$L_A = \frac{(\sum_{v=1}^{5 \times 5} 1)^2}{\sum_{v=1}^{5 \times 5} 1^2} = \frac{625}{25} = 25$$

$$L_A = \frac{(\sum_{v=1}^{3 \times 5} 1)^2}{\sum_{v=1}^{3 \times 5} 1^2} = \frac{225}{15} = 15$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{40} + \frac{1}{25} + \frac{1}{15}} = 22.785$$

In Table 13 each of the generalizability coefficients (G and D) are calculated as a ratio of: true variance/(true variance + error variance). “True variance” is signified by the term “tau” (and in general is equal to the variance of the facet of differentiation, more below) and the denominator represents “tau” + error variance. Notice that for the D-coefficients, more terms are added into the denominator that will result in lower values for D versus G coefficients. Error variances are comprised as indicated in the notes after Table 13.

For nested facets of differentiation, such as patients nested in doctors, it is imperative that the researcher understands the true variance that they seek to measure. According to Brennan's (2001a) discussion on group means (5.3, page 157) the generalizability coefficient (G) of the facet of differentiation (p:d) can assess the generalizability of the patient ratings for “a randomly selected doctor” and would result in the formula $G(p:d) = \sigma^2(p:d)/[\sigma^2(p:d) + \sigma^2(p_i:d)/\text{levels}_{p_i:d}]$. Where tau represents the variance component of p:d (0.092) and the error variance is 0.157/5. If, however, we wanted the generalizability coefficient that represents patient ratings over “all doctors” then the tau needs to include both the p:d and d variance components [(tau = $\sigma^2(p:d) + \sigma^2(d)$), 0.092 and .002, respectively]. We will assume, in this example exercise, that we want the latter interpretation of generalizability for patient ratings.

Table 13. Calculating G and D Coefficients for Narayanan et al. (2010) data

| Generalizability Coefficient | Facet of Differentiation = P:D | Facet of Differentiation = I | Facet of Differentiation = D |
|------------------------------|---|--|--|
| Generalizability (G) | $G = \text{Tau}/(\text{Tau} + \text{EV}(G)^*)$ $\text{Tau} = \text{VC}^{**}(\text{P:D}) + \text{VC}(\text{D})$ $\text{Tau} = .092 = .092$ $\text{EV}(G) = \text{VC}(\text{P:I:D})/\text{Level_P:D}(\text{P:I:D}) = .1572/5 = .031$ $G = .092/((.092+.031))$ G = .752 | $G = \text{Tau}/(\text{Tau} + \text{EV}(G))$ $\text{Tau} = \text{VC}(\text{I})$ $\text{Tau} = .028$ $\text{EV}(G) = \text{VC}(\text{IxD})/\text{Level_I}(\text{IxD}) + \text{VC}(\text{P:I:D})/\text{Level_I}(\text{P:I:D}) = 0/2.61 + .157/16 = .010$ $G = .028/((.028+.010))$ G = .737 | $G = \text{Tau}/(\text{Tau} + \text{EV}(G))$ $\text{Tau} = \text{VC}(\text{D})$ $\text{Tau} = .002$ $\text{EV}(G) = \text{VC}(\text{P:D})/\text{Level_D}(\text{P:D}) + \text{VC}(\text{IxD})/\text{Level_D}(\text{IxD}) + \text{VC}(\text{P:I:D})/\text{Level_D}(\text{P:I:D}) = .092/4.557+0/5 + .157/22.785 = .027$ $G = .002/((.002+.027))$ G = .069 |
| | $D = \text{Tau}/(\text{Tau} + \text{EV}(D)^{***})$ | $D = \text{Tau}/(\text{Tau} + \text{EV}(D))$ | $D = \text{Tau}/(\text{Tau} + \text{EV}(D))$ |

| | | | |
|--------------------------|---|---|--|
| Dependability (D) | $EV(D) = (.028/5) + (0/5)^{****} + (.157/5)$ $= .037$ $D = .094/ (.094 + .037)$ D = .718 | $EV(D) = (.002/2.61) + (.092/16) + (0/2.61) + (.157/16)$ $= .016$ $D = .028/ (.028 + .016)$ D = .636 | $EV(D) = (.028/5) + (.092/4.557) + (0/5) + (.157/22.785)$ $= .033$ $D = .002/ (.002 + .033)$ D = .057 |
|--------------------------|---|---|--|

*EV(G) = Respective Error Variance for G = sum of all VCs (except the VC of differentiation) that contain the facet of differentiation and at least one facet of generalization, each divided by the size of the product of the respective facets of generalization

**VC = Respective Variance Component for facet of differentiation

*** EV(D) = Respective Error Variance for D = sum of all VCs (except the VC of differentiation), each divided by the size of the product of the respective facets of generalization

****Because the DI variance component term is negative (-.016), we will substitute “0” for it when calculating the error variances for G and D, as suggested by Brennan (2001a).

It is worth pointing out at this stage when we finally obtain our Generalizability coefficient (G) for the P:D effect (0.752), that is it substantially different from that reported in the Narayanan et al. (2010) paper (0.91 for “unaggregated G” and 0.80 for “aggregated G”, p. 371). This is due to the completely different approach used. We will continue to observe the more traditional approach introduced by Henderson and made popular by Brennan (2001a).

Decision Study (D-Study) Coefficients

Decision Studies utilize the variance components calculated from Generalizability Studies and manipulate the levels of facets to show how generalizability coefficients may be improved. For example, in this study we saw that the generalizability coefficient for p:d was below 0.80, and a researcher may wonder how the coefficient would change if there were more items for the same variance. With balanced and complete studies this is straightforward as the user can just input a dictionary of potential study combinations. For the items x (patients:doctors) this may look like:

D_Study_Dict: {items: [6, 7, 8], patients: [5], doctors: [3]}

Such a D-Study would calculate G Coefficients for the balanced scenarios where there are 3 unique doctors with 5 unique patients per doctor, each answering 6, 7, or 8 items per study respectively. Working out the 6-item scenario, we would have the following table of pseudo counts:

Table 14: Pseudo Counts Table for Unbalanced D Study (6 items, 3 Doctors, 5 patients)

| Doctor | Patient | Item |
|--------|---------|------|
| A | 1 | 1 |
| A | 1 | 2 |
| A | 1 | 3 |
| A | 1 | 4 |
| A | 1 | 5 |
| A | 1 | 6 |

... (table con't)

| Doctor | Patient | Item |
|--------|---------|------|
| A | 5 | 1 |
| A | 5 | 2 |
| A | 5 | 3 |
| A | 5 | 4 |
| A | 5 | 5 |
| A | 5 | 6 |

... (table con't)

| Doctor | Patient | Item |
|--------|---------|------|
| B | 10 | 1 |
| B | 10 | 2 |
| B | 10 | 3 |
| B | 10 | 4 |

| | | |
|----------|-----------|----------|
| B | 10 | 5 |
| B | 10 | 6 |

... (table con't)

| Doctor | Patient | Item |
|---------------|----------------|-------------|
| C | 15 | 1 |
| C | 15 | 2 |
| C | 15 | 3 |
| C | 15 | 4 |
| C | 15 | 5 |
| C | 15 | 6 |
| | | |

We then would have new levels coefficients to evaluate p:d (Table 15):

Table 15. Modified Levels for the D-Study {items: 6, patients: 5, doctors: 3}

| Variance Term | P:D Facet of Differentiation Denominator for G and D |
|----------------------------|---|
| $\sigma^2(d)$ | 1 |
| $\sigma^2(p:d)$ | 1 |
| $\sigma^2(i)$ | 6 |
| $\sigma^2(i \times d)$ | 6 |
| $\sigma^2(i \times (p:d))$ | 6 |

With the adjusted generalizability coefficients (Table 16):

Table 16. G Coefficients for the D-Study {items: 6, patients: 5, doctors: 3}:

| Generalizability Coefficient | Facet of Differentiation = P:D |
|-------------------------------------|--|
| Generalizability (G) | $G = \text{Tau} / (\text{Tau} + \text{EV}(G)^*)$ $\text{Tau} = \text{VC}^{**}(\text{P:D}) + \text{VC}(\text{D})$ $\text{Tau} = .092 + .002 = .094$ $\text{EV}(G) = \text{VC}(\text{Pl:D}) / \text{Level_P:D}(\text{Pl:D})$ $= .1572 / 6 = .026$ $G = .094 / (.094 + .026)$ G = .78 |
| Dependability (D) | $D = \text{Tau} / (\text{Tau} + \text{EV}(D)^{***})$ $\text{EV}(D) = (.028 / 6) + (0 / 5)^{****} + (.157 / 6)$ $= .031$ $D = .094 / (.094 + .031)$ D = .752 |

We see in this example that the G and D coefficients for P:D improve with the addition of another item (and assuming that

there are 5 patients rating 3 doctors), indicating that for the same conditions this study would have been more generalizable had a greater number of items been included.

However, D-Studies *need not be balanced*. Let's turn our attention to the generalizability coefficients associated with the "items" facet (which indicates the effect of questionnaire items in delineating doctor quality from patients) if some other aspect of the study was modified. In the G-Study, recall we had 5 items and the following unbalanced nesting of patients under doctors: {A: 8, B:5, C:3}. A researcher may have 16 patients but wonder: "If the number of patients nested under doctors A, B, and C were 6, 5, 5 respectively, how would this affect the generalizability of "items"? For demonstration let's perform a D Study with the following characteristics and associated pseudo counts table (Table 17):

D-Study: {items: 5, {A: 6, B: 5, C, 5}}

Table 17. Pseudo Counts for D-Study: {items: 5, {A: 6, B: 5, C, 5}}

| Doctor | Patient | Item |
|--------|---------|------|
| A | 1 | 1 |
| A | 1 | 2 |
| A | 1 | 3 |
| A | 1 | 4 |
| A | 1 | 5 |

... (table con't)

| Doctor | Patient | Item |
|--------|---------|------|
| A | 6 | 1 |
| A | 6 | 2 |
| A | 6 | 3 |
| A | 6 | 4 |
| A | 6 | 5 |

... (table con't)

| Doctor | Patient | Item |
|--------|---------|------|
| B | 11 | 1 |
| B | 11 | 2 |
| B | 11 | 3 |
| B | 11 | 4 |
| B | 11 | 5 |

... (table con't)

| Doctor | Patient | Item |
|--------|---------|------|
| C | 16 | 1 |
| C | 16 | 2 |
| C | 16 | 3 |
| C | 16 | 4 |
| C | 16 | 5 |

Using the method to calculate levels presented in Table 12, these new pseudo counts would result in the following levels (Table 18) for facet of differentiation, item.

Table 18. Modified Levels for the D-Study: {items: 5, {A: 6, B: 5, C, 5}}

| Variance Term | I Facet of Differentiation Denominator for G and D |
|----------------------------|--|
| $\sigma^2(d)$ | 2.98 |
| $\sigma^2(p:d)$ | 16 |
| $\sigma^2(i)$ | 1 |
| $\sigma^2(i \times d)$ | 2.98 |
| $\sigma^2(i \times (p:d))$ | 16 |

$\sigma^2(d)$ and $\sigma^2(i \times (p:d))$ remain the same because there are still 16 unique combinations of p and d despite being distributed differently. However, $\sigma^2(d)$ and $\sigma^2(i \times d)$ change because of this distribution. Example calculation for $\sigma^2(d)$ shown below (which is identical to $\sigma^2(i \times d)$):

$C(i, d) = |p_v|$, for $p = [6, 5, 5]$ with $v \in [1, 3]$; $g \in [1, 5]$. Since i is not involved with the unbalanced nesting, all levels of i , $i \in [1, 5]$, will be the same.

$$L = L_{1-5} = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(6+5+5)^2}{(6^2+5^2+5^2)} = \frac{256}{86} = 2.98$$

As a small aside, the reader should notice that as we have nearly balanced data (6, 5, 5) the levels of $\sigma^2(d)$ on facet of differentiation, i , are nearly equivalent to the balanced case ($L=3$ for 3 doctors).

Using our modified levels coefficients we can calculate the D-Study coefficients as follows (Table 19):

Table 19. G Coefficients for the D-Study: {items: 5, {A: 6, B: 5, C, 5}}

| Facet of Differentiation = I |
|---|
| $G = \text{Tau}/(\text{Tau} + \text{EV}(G))$ $\text{Tau} = \text{VC}(I)$ $\text{Tau} = .028$ $\text{EV}(G) = \text{VC}(I \times D)/\text{Level}_I(I \times D)$ $+ \text{VC}(P \times I \times D)/\text{Level}_I(P \times I \times D)$ $= 0/2.98 + .157/16$ $= .010$ $G = .028/ (.028 + .010)$ G = .737 |
| $D = \text{Tau}/ (\text{Tau} + \text{EV}(D))$ $\text{EV}(D) = (.002/2.98) +$ $(.092/16) + (0/2.98) +$ $(.157/16)$ $= .016$ $D = .028/ (.028 + .016)$ D = .636 |

Our generalizability coefficients, G and D remain the same as in the prior example! This is a direct result of the fact that the D and DI variances are so low that they hardly affect the calculation (in the first example or this one), thus for any reasonable choice of 16 $p:d$ these coefficients will remain approximately the same. Of course, the numbers of P and D may be increased to reduce these variances and improve the generalizability coefficients, but those exercises will be left for the reader.

Additional notes Regarding urGENOVA and G-String_V

urGENOVA and G-String_V programs also calculate the Estimated Mean Squares and from those the estimated Sums of Squares. However, neither of these are needed for calculating G and D coefficients. In addition, as Brennan (2001a) notes, because the Estimated Mean Square equations are complicated, expressions for estimators of the variance components in terms of mean square are also complicated. It is simpler to express the estimators of the variance components with respect to T-terms.

Another Sample Data Set

Another data set (the Narayana et al., 2010 with changed values for Doctor B (all ratings for all items increased by 1 point) and Doctor C (all ratings for all items lowered by 1 point) was generated (Table 20). This, not surprisingly, dramatically increases the variance on the “Doctor” facet, and the subsequent generalizability coefficients on which they are based.

Table 20. Modified Narayanan data set for more variance on the Doctor variable

| Doctor | Patient | item1 | item2 | item3 | item4 | item5 |
|--------|---------|-------|-------|-------|-------|-------|
| A | 1 | 4 | 4 | 4 | 4 | 4 |
| A | 2 | 4 | 4 | 3 | 3 | 4 |
| A | 3 | 3 | 4 | 4 | 3 | 4 |
| A | 4 | 3 | 3 | 3 | 3 | 3 |
| A | 5 | 3 | 3 | 4 | 4 | 3 |
| A | 6 | 4 | 4 | 3 | 3 | 3 |
| A | 7 | 3 | 3 | 3 | 3 | 3 |
| A | 8 | 4 | 4 | 3 | 3 | 3 |
| B | 9 | 5 | 5 | 5 | 5 | 4 |
| B | 10 | 5 | 5 | 5 | 5 | 5 |
| B | 11 | 5 | 5 | 5 | 5 | 5 |
| B | 12 | 5 | 4 | 4 | 4 | 4 |
| B | 13 | 5 | 5 | 4 | 4 | 4 |
| C | 14 | 3 | 3 | 2 | 2 | 2 |
| C | 15 | 2 | 2 | 2 | 2 | 2 |
| C | 16 | 3 | 3 | 3 | 3 | 2 |

The resulting Table 21 shows that all the T-values change (as they would be expected to do), but not the “count”-based values for the coefficients, since these are only based on the count-relevant values for each coefficient.

Table 21. Facet and Variance Coefficients for modified Narayanan data set

| Facet | $\sigma^2(d)$ | $\sigma^2(p:d)$ | $\sigma^2(i)$ | $\sigma^2(di)$ | $\sigma^2(pi:d)$ | μ^2 | T-Value |
|----------------------------|---------------|-----------------|---------------|----------------|------------------|---------|----------|
| D | 80 | 15 | 16 | 16 | 3 | 80 | 1103.185 |
| P:D | 80 | 80 | 16 | 16 | 16 | 80 | 1111.200 |
| I | 30.625 | 5 | 80 | 30.625 | 5 | 80 | 1053.250 |
| DI | 80 | 15 | 80 | 80 | 15 | 80 | 1105.808 |
| PI:D (total for Henderson) | 80 | 80 | 80 | 80 | 80 | 80 | 1122.000 |
| Mean | 30.625 | 5 | 16 | 6.125 | 1 | 80 | 1051.250 |

The resulting Table 22 of variance components demonstrates that while the D variance component is dramatically changed, the others remain exactly the same.

Table 22. Variance Components from modified Narayanan et al. (2010) data

| Facet | Variance Component | Proportion of Variance |
|-------|--------------------|------------------------|
| D | 1.030 | .788 |
| P:D | .092 | .070 |
| I | .028 | .021 |
| DI | -.016 | 0 |
| PI:D | .157 | .120 |

This is a much more easily-interpretable set of variance components: Most of the variance is due to different ratings of the doctors (79%). That is, the doctors are performing differently as rated by their patients. The P:D effect is 7%, indicating that the patients are very consistent (not much variance is ascribed to them) in rating the different doctors. The items are also internally consistent with only 2% of the variance in the data set ascribed to using different items. The resulting generalizability coefficients for the P:D and D facets goes up dramatically (Table 23):

Table 23. G and D Coefficients for modified Narayanan et al. (2010) data

| Generalizability Coefficient | Facet of Differentiation = P:D | Facet of Differentiation = I | Facet of Differentiation = D |
|------------------------------|--------------------------------|------------------------------|------------------------------|
|------------------------------|--------------------------------|------------------------------|------------------------------|

| | | | |
|-----------------------------|---|--|---|
| Generalizability (G) | $G = \text{Tau}/(\text{Tau} + \text{EV(G)})$ $\text{Tau} = \text{VC(P:D)} + \text{VC(D)}$ $\text{Tau} = .092 + 1.030 = 1.122$ $\text{EV(G)} = \text{VC(PI:D)}/\text{Level_P:D(PI:D)}$ $= .157/5 = .031$ $G = 1.122/(1.122 + .031)$ G = .973 | $G = \text{Tau}/(\text{Tau} + \text{EV(G)})$ $\text{Tau} = \text{VC(I)}$ $\text{Tau} = .028$ $\text{EV(G)} = \text{VC(IxD)}/\text{Level_I(IxD)} + \text{VC(PI:D)}/\text{Level_I(PI:D)}$ $= 0/2.61 + .157/16 = .010$ $G = .028/(.028 + .010)$ G = .737 | $G = \text{Tau}/(\text{Tau} + \text{EV(G)})$ $\text{Tau} = \text{VC(D)}$ $\text{Tau} = 1.030$ $\text{EV(G)} = \text{VC(P:D)}/\text{Level_D(P:D)} + \text{VC(IxD)}/\text{Level_D(IxD)} + \text{VC(PI:D)}/\text{Level_D(PI:D)}$ $= .092/4.557 + 0/5 + .157/22.785 = .027$ $G = 1.030/(1.030 + .027)$ G = .974 |
| Dependability (D) | $D = \text{Tau}/(\text{Tau} + \text{EV(D)})$ $\text{EV(D)} = (.0276/5) + (0/5) + (.1572/5) = .037$ $D = 1.122/(1.122 + .037)$ D = .968 | $D = \text{Tau}/(\text{Tau} + \text{EV(D)})$ $\text{EV(D)} = (1.030/2.61) + (.092/16) + (0/2.61) + (.157/16) = .410$ $D = .028/(.028 + .410)$ D = .064 | $D = \text{Tau}/(\text{Tau} + \text{EV(D)})$ $\text{EV(D)} = (.028/5) + (.092/4.557) + (0/5) + (.157/22.785) = .033$ $D = 1.030/(1.030 + .033)$ D = .969 |

Missing Data

Up to this point, there have been no missing data points in the calculations. However, many data sets do have missing data (in fact Henderson's 1953 data set is mostly missing data). The procedure for calculating the variance components and denominators for calculating G and D values is presented next. **Note that NO Decision-study follow ups can be done with missing data sets.** We will use Narayanan et al.'s original (2010) data set again to go through the example with the following changes: There is 1 data point missing for each "doctor" (highlighted in red) (see Table 24). Note that the "counts" of ratings that go into the denominator of the "Squared sums" VARIES as opposed to remaining constant across the facet.

Table 24. Narayanan et al. (2010) data set with 3 missing data points.

| Doctor | Patient | item1 | item2 | item3 | item4 | item5 | Sum across Patient Ratings | Squared sum across Patient Ratings | Squared sum across patient ratings/#ratings each completed |
|--------|--------------------------|---------|---------|---------|-------|-------|----------------------------|------------------------------------|--|
| A | 1 | missing | 4 | 4 | 4 | 4 | 16 | 256 | 64 |
| A | 2 | 4 | 4 | 3 | 3 | 4 | 18 | 324 | 64.8 |
| A | 3 | 3 | 4 | 4 | 3 | 4 | 18 | 324 | 64.8 |
| A | 4 | 3 | 3 | 3 | 3 | 3 | 15 | 225 | 45 |
| A | 5 | 3 | 3 | 4 | 4 | 3 | 17 | 289 | 57.8 |
| A | 6 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| A | 7 | 3 | 3 | 3 | 3 | 3 | 15 | 225 | 45 |
| A | 8 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| B | 9 | 4 | 4 | 4 | 4 | 3 | 19 | 361 | 72.2 |
| B | 10 | 4 | missing | 4 | 4 | 4 | 16 | 256 | 64 |
| B | 11 | 4 | 4 | 4 | 4 | 4 | 20 | 400 | 80 |
| B | 12 | 4 | 3 | 3 | 3 | 3 | 16 | 256 | 51.2 |
| B | 13 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| C | 14 | 4 | 4 | 3 | 3 | 3 | 17 | 289 | 57.8 |
| C | 15 | 3 | 3 | missing | 3 | 3 | 12 | 144 | 36 |
| C | 16 | 4 | 4 | 4 | 4 | 3 | 19 | 361 | 72.2 |
| | Item Rating Sums | 55 | 55 | 52 | 54 | 53 | | | |
| | Item Rating sums squared | 3025 | 3025 | 2704 | 2916 | 2809 | | | |

| | | | | | | | | | |
|--|---|---------|----------|----------|--------|----------|--|--|--|
| | Item sums squared/#ratings For each item | 201.667 | 201.6667 | 180.2667 | 182.25 | 175.5625 | | | |
|--|---|---------|----------|----------|--------|----------|--|--|--|

Next we calculate the T-Values for each facet:

1. Doctor (D)

Sum all ratings across each item for each Patient; Sum the individual patient ratings for each doctor; square these sums; divide by the number of rating counts within each doctor; add the quotients.

$$\text{Doctor A} = ((16+18+18+15+17+17+15+17)^2)/39 = 453.564$$

$$\text{Doctor B} = ((19+16+20+16+17)^2)/24 = 322.667$$

$$\text{Doctor C} = ((17+12+19)^2)/14 = 164.571$$

$$\text{Sum across all 3 quotients} = 940.802$$

2. Patient:Doctor (P:D)

Sum all ratings across each item for each Patient; Square these sums; divide by the number of rating counts within each patient; sum across these quotients.

$$\text{p:d 1 } (16^2)/4 = 256/4 = 64$$

$$\text{p:d 2 } (18^2)/5 = 324/5 = 64.8$$

.....

$$\text{p:d 15 } (12^2)/4 = 144/4 = 36$$

$$\text{p:d 16 } (19^2)/5 = 362; 361/5 = 72.2$$

$$\text{Sum across all 16 quotients} = 948.2$$

3. Item (I)

Sum all ratings down each item; Square these sums; divide by the number of rating counts within each item; sum across these quotients.

$$\text{Item 1 } (55^2)/15 = 201.667$$

$$\text{Item 2 } (55^2)/15 = 201.667$$

$$\text{Item 3 } (52^2)/15 = 180.267$$

$$\text{Item 4 } (54^2)/16 = 182.250$$

$$\text{Item 5 } (53^2)/16 = 175.563$$

$$\text{Sum across all 5 quotients} = 941.413$$

4. Doctor X Item (DI)

Sum down each set of items for each doctor (there will be 15 combinations); square these sums; divide each sum by the number of ratings that go into each of the 15 combinations; sum the quotients.

$$\text{di1} = (4+3+3+3+4+3+4) = 24; 24^2 = 576; 576/7 = 82.286$$

$$\text{di2} = (4+4+4+3+3+4+3+4) = 29; 29^2 = 841; 841/8 = 105.125$$

.....

$$\text{di6} = (4+4+4+4+4) = 20; 20^2 = 400; 400/5 = 80.000$$

$$\text{di7} = (4+4+3+4) = 15; 15^2 = 225; 361/4 = 56.250$$

.....

$$\text{di14} = (3+3+4) = 10; 10^2 = 100; 100/3 = 33.333$$

$$\text{di15} = (3+3+3) = 9; 9^2 = 81; 81/3 = 27.000$$

$$\text{Sum across all quotients} = 943.311$$

5. PatientXItem:Doctor (PI:D)

Note that this is the total uncorrected sums of squares. Each individual rating is first squared. Then these are summed across all ratings.

Going across each row....

$$0^2 + 4^2 + 4^2 + 4^2 + 4^2 +$$

$$4^2 + 4^2 + 4^2 + 4^2 + 3^2 +$$

....continuing for all rows

$$3^2 + 3^2 + 0^2 + 3^2 + 3^2 +$$

$$4^2 + 4^2 + 4^2 + 4^2 + 3^2$$

Summing across all values = 959.00

6. Mean

Sum FIRST across all values; square this sum; divide by the total number of ratings that went into the sum.

$$0 + 4 + 4 + 4 + 4 +$$

$$4 + 4 + 4 + 4 + 3 +$$

....continuing for all rows

$$3 + 3 + 0 + 3 + 3 +$$

$$4 + 4 + 4 + 4 + 3$$

$$\text{Summing across all values} = (269^2)/77 = 939.753$$

We can now put our Facet T-values into Table 25. below.

Table 25. Facets and T-values for the Narayanan data set

| Facet | T-Value |
|--|---------|
| D (doctor) | 940.802 |
| P:D (patient nested in doctor) | 948.200 |
| I (item) | 941.413 |
| DI (doctor X item) | 943.311 |
| PI:D (patient X item nested in doctor) | 959.000 |
| Mean | 939.753 |

As always, obtaining the “counts” that go into the analysis of the variance components, is tedious. It becomes even more tedious when there are missing values. Let’s begin....

As before, we first put in the “easy” ones.

1. For the μ^2 term, ALL coefficients are based on the total number of data points involved in the study (denoted N; 77 for this data set).
2. For the PI:D (Total which includes the highest level term plus error) facet, all coefficients are also based on the total number of data points involved in the study (N).
3. For each facet (except the Mean), the coefficient for that variance is also the total number of data points involved in the study (N). For example, the coefficient for $\sigma^2(d)$ for effect D is N.
4. For the $\sigma^2(\pi:d)$ term column, the sample sizes are equal to the number of levels of that facet (D = 3, P:D = 16, I = 5, DI = 15 (there are 15 different DI combinations), for the Mean is always = 1, and for the highest order effect (PI:D – also known as the total by Henderson) is always = N.

This leaves 16 “?” values to “fill in” (Table 26).

Table 26. “Counts” for use in calculating the Variance Components with Narayanan et al. (2010) with missing data

| effect | counts for effects | | | | | | T-values |
|--------|--------------------|------------|----------|-----------|-------------|-----------|----------|
| | counts D | counts P:D | counts I | counts DI | counts pi:d | counts mu | |
| D | 77 | ? | ? | ? | 3 | 77 | 940.802 |
| P:D | ? | 77 | ? | ? | 16 | 77 | 948.200 |
| I | ? | ? | 77 | ? | 5 | 77 | 941.413 |
| DI | ? | ? | ? | 77 | 15 | 77 | 943.311 |
| pi:D,e | 77 | 77 | 77 | 77 | 77 | 77 | 959.000 |
| Mean | ? | ? | ? | ? | 1 | 77 | 939.753 |

1. The $\sigma^2(p:d)$ term on D (data are collapsed across Items). Sample size starts by calculating the squared sum of:
counts of each p:d term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((4^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2) + (5^2))/39 = 191/39 = 4.897 +$$

$$D2: ((5^2) + (4^2) + (5^2) + (5^2) + (5^2))/24 = 116/24 = 4.833 +$$

$$D3: ((5^2) + (4^2) + (5^2))/14 = 66/14 = 4.714$$

$$\text{Sum} = 14.444$$

2. The $\sigma^2(i)$ term on D (data are collapsed across P:D). Sample size starts by calculating the squared sum of:
counts of each i term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((7^2) + (8^2) + (8^2) + (8^2) + (8^2))/39 = 305/39 = 7.820 +$$

$$D2: ((5^2) + (4^2) + (5^2) + (5^2) + (5^2))/24 = 116/24 = 4.833 +$$

$$D3: ((3^2) + (3^2) + (2^2) + (3^2) + (3^2))/14 = 40/14 = 2.857$$

$$\text{Sum} = 15.510$$

3. The $\sigma^2(di)$ term on D (data are collapsed across P:D). Sample size starts by calculating the squared sum of:
counts of each di term within each D, divided by the total number of counts for that D. Then sum the quotients.

$$D1: ((7^2) + (8^2) + (8^2) + (8^2) + (8^2))/39 = 305/39 = 7.820 +$$

$$D2: ((5^2) + (5^2) + (5^2) + (5^2) + (5^2))/24 = 116/24 = 4.833 +$$

$$D3: ((3^2) + (3^2) + (3^2) + (3^2) + (3^2))/14 = 40/14 = 2.857$$

$$\text{Sum} = 15.510$$

4. The $\sigma^2(d)$ term on P:D (data are collapsed across Items). Sample size starts by calculating the squared sum of:
counts of each d term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

$$PD\ 1: (4^2)/4 = 4 +$$

$$PD\ 2: (5^2)/5 = 5 +$$

$$PD\ 3: (5^2)/5 = 5 +$$

$$PD\ 4: (5^2)/5 = 5 +$$

$$PD\ 5: (5^2)/5 = 5 +$$

$$PD\ 6: (5^2)/5 = 5 +$$

$$PD\ 7: (5^2)/5 = 5 +$$

$$PD\ 8: (5^2)/5 = 5 +$$

$$PD\ 9: (5^2)/5 = 5 +$$

$$PD\ 10: (4^2)/4 = 4 +$$

$$PD\ 11: (5^2)/5 = 5 +$$

$$PD\ 12: (5^2)/5 = 5 +$$

$$PD\ 13: (5^2)/5 = 5 +$$

$$PD\ 14: (5^2)/5 = 5 +$$

$$PD\ 15: (4^2)/4 = 4 +$$

$$PD\ 16: (5^2)/5 = 5$$

$$\text{Sum} = 77$$

5. The $\sigma^2(i)$ term on P:D. Sample size starts by calculating the squared sum of:
counts of each i term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

$$PD\ 1: ((1^2) + (1^2) + (1^2) + (1^2))/4 = 1 +$$

$$PD\ 2: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

$$PD\ 3: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

$$PD\ 4: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

PD 5: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 6: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 7: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 8: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 9: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 10: $((1^2) + (1^2) + 1^2 + (1^2))/4 = 1 +$
 PD 11: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 12: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 13: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 14: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 15: $((1^2) + (1^2) + 1^2 + (1^2))/4 = 1 +$
 PD 16: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1$

Sum = 16

6. The $\sigma^2(d_i)$ term on P:D. Sample size starts by calculating the squared sum of:
 # counts of each d_i term within each P:D, divided by the total number of counts for that P:D. Then sum the quotients.

PD 1: $((1^2) + (1^2) + 1^2 + (1^2))/4 = 1 +$
 PD 2: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 3: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 4: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 5: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 6: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 7: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 8: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 9: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 10: $((1^2) + (1^2) + 1^2 + (1^2))/4 = 1 +$
 PD 11: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 12: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 13: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 14: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1 +$
 PD 15: $((1^2) + (1^2) + 1^2 + (1^2))/4 = 1 +$
 PD 16: $((1^2) + (1^2) + (1^2) + 1^2 + (1^2))/5 = 1$

Sum = 16

7. The $\sigma^2(d)$ term on I. Sample size starts by calculating the squared sum of:
 # counts of each d term within each I, divided by the total number of counts for that I. Then sum the quotients.

I 1: $((7^2) + (5^2) + (3^2))/15 = 83/15 = 5.533 +$
 I 2: $((8^2) + (4^2) + (3^2))/15 = 89/15 = 5.933 +$
 I 3: $((8^2) + (5^2) + (2^2))/15 = 93/15 = 6.200 +$
 I 4: $((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 +$
 I 5: $((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125$

Sum = 29.916

8. The $\sigma^2(p:d)$ term on I. Sample size starts by calculating the squared sum of:
 # counts of each $p:d$ term within each I, divided by the total number of counts for that I. Then sum the quotients.

I 1: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/15 = 15/15 = 1 +$
 I 2: $((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/15 = 15/15 = 1 +$

$$I 3: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/15 = 15/15 = 1 +$$

$$I 4: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/16 = 16/16 = 1 +$$

$$I 5: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/16 = 16/16 = 1$$

$$\text{Sum} = 5$$

9. The $\sigma^2(d_i)$ term on I. Sample size starts by calculating the squared sum of:

counts of each d_i term within each I, divided by the total number of counts for that I. Then sum the quotients.

$$I 1: ((7^2) + (5^2) + (3^2))/15 = 83/15 = 5.533 +$$

$$I 2: ((8^2) + (4^2) + (3^2))/15 = 89/15 = 5.933 +$$

$$I 3: ((8^2) + (5^2) + (2^2))/15 = 93/15 = 6.200 +$$

$$I 4: ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125 +$$

$$I 5: ((8^2) + (5^2) + (3^2))/16 = 98/16 = 6.125$$

$$\text{Sum} = 29.916$$

10. The $\sigma^2(d)$ term on DI. Sample size starts by calculating the squared sum of:

counts of each d term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

$$DI 1: (7^2)/7 = 7 +$$

$$DI 2: (8^2)/8 = 8 +$$

$$DI 3: (8^2)/8 = 8 +$$

$$DI 4: (8^2)/8 = 8 +$$

$$DI 5: (8^2)/8 = 8 +$$

$$DI 6: (5^2)/5 = 5 +$$

$$DI 7: (4^2)/4 = 4 +$$

$$DI 8: (5^2)/5 = 5 +$$

$$DI 9: (5^2)/5 = 5 +$$

$$DI 10: (5^2)/5 = 5 +$$

$$DI 11: (3^2)/3 = 3 +$$

$$DI 12: (3^2)/3 = 3 +$$

$$DI 13: (2^2)/2 = 2 +$$

$$DI 14: (3^2)/3 = 3 +$$

$$DI 15: (3^2)/3 = 3$$

$$\text{Sum} = 77$$

11. The $\sigma^2(p:d)$ term on DI. Sample size starts by calculating the squared sum of:

counts of each $p:d$ term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

$$DI 1: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/7 = 1 +$$

$$DI 2: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$$

$$DI 3: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$$

$$DI 4: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$$

$$DI 5: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2) + (1^2))/8 = 1 +$$

$$DI 6: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

$$DI 7: ((1^2) + (1^2) + (1^2) + (1^2))/4 = 1 +$$

$$DI 8: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

$$DI 9: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

$$DI 10: ((1^2) + (1^2) + (1^2) + (1^2) + (1^2))/5 = 1 +$$

DI 11: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 12: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 13: $((1^2) + (1^2))/2 = 1 +$
 DI 14: $((1^2) + (1^2) + (1^2))/3 = 1 +$
 DI 15: $((1^2) + (1^2) + (1^2))/3 = 1$
Sum = 15

12. The $\sigma^2(i)$ term on DI. Sample size starts by calculating the squared sum of:
 # counts of each i term within each DI, divided by the total number of counts for that DI. Then sum the quotients.

DI 1: $(7^2)/7 = 7 +$
 DI 2: $(8^2)/8 = 8 +$
 DI 3: $(8^2)/8 = 8 +$
 DI 4: $(8^2)/8 = 8 +$
 DI 5: $(8^2)/8 = 8 +$
 DI 6: $(5^2)/5 = 5 +$
 DI 7: $(4^2)/4 = 4 +$
 DI 8: $(5^2)/5 = 5 +$
 DI 9: $(5^2)/5 = 5 +$
 DI 10: $(5^2)/5 = 5 +$
 DI 11: $(3^2)/3 = 3 +$
 DI 12: $(3^2)/3 = 3 +$
 DI 13: $(2^2)/2 = 2 +$
 DI 14: $(3^2)/3 = 3 +$
 DI 15: $(3^2)/3 = 3$
Sum = 77

13. The $\sigma^2(d)$ term on mean. Sample size starts by calculating the squared sum of:
 # counts of each d term within each mean, divided by the total number of counts for the mean.

Mean: $((39^2) + (24^2) + (14^2))/77 = (1521 + 576 + 196)/77 = 2293/77 = 29.779$

14. The $\sigma^2(p:d)$ term on mean. Sample size starts by calculating the squared sum of:
 # counts of each p:d term within each mean, divided by the total number of counts for the mean.

Mean: $((4^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(5^2)+(4^2)+(5^2)+(5^2)+(5^2)+(5^2)+(4^2)+(5^2))/77 = ((16*3)+(25*13))/77 = (48+325)/77 = 373/77 = 4.844$

15. The $\sigma^2(i)$ term on mean. Sample size starts by calculating the squared sum of:
 # counts of each i term within each mean, divided by the total number of counts for the mean

Mean: $((15^2) + (15^2) + (15^2) + (16^2) + (16^2))/77 = (225*3) + (256*2))/77 = (675+512)/77 = 1187/77 = 15.416$

16. The $\sigma^2(di)$ term on mean. Sample size starts by calculating the squared sum of:
 # counts of each di term within each mean, divided by the total number of counts for the mean.

Mean: $((7^2) + (5^2)+(3^2)+(8^2)+(4^2)+(3^2)+(8^2)+(5^2)+(2^2)+(8^2)+(5^2)+(3^2)+(8^2)+(5^2)+(3^2))/77 = ((64*4) + (25*4) + (9*4)+49+16+4)/77 = (256+100+36+69)/77 = 461/77 = 5.987$

We can now insert the calculated values into our “?” cells (Table 27) and using matrix equations or regression, solve for the Variance Components.

| effect | counts for effects | | | | | | | variance component |
|--------|--------------------|------------|----------|-----------|-------------|-----------|----------|--------------------|
| | counts D | counts P:D | counts I | counts DI | counts pi:d | counts mu | T-values | |
| D | 77 | 14.444 | 15.510 | 15.510 | 3 | 77 | 940.802 | 0.001 |
| P:D | 77 | 77 | 16 | 16 | 16 | 77 | 948.200 | 0.083 |
| I | 29.916 | 5 | 77 | 29.916 | 5 | 77 | 941.413 | 0.021 |
| DI | 77 | 15 | 77 | 77 | 15 | 77 | 943.311 | -0.014 |
| pi:D,e | 77 | 77 | 77 | 77 | 77 | 77 | 959.000 | 0.170 |
| Mean | 29.779 | 4.844 | 15.416 | 5.987 | 1 | 77 | 939.753 | 12.1936 |

33

$$L_g = \frac{(\sum_{v=1}^5 1)^2}{\sum_{v=1}^5 1^2} = 5$$

$$L_{g'} = \frac{(\sum_{v=1}^4 1)^2}{\sum_{v=1}^4 1^2} = 4$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{16}{\frac{1}{4} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{4} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{4} + \frac{1}{5}} = 4.78$$

6. Level of the $\sigma^2(d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, d) = |p_v|$, for $p = [8^*, 5^*, 3^*]$ with $v \in [1, 3]$; $g \in [1, 5]$ and *indicating missing data. Since i is not involved with the unbalanced nesting, all levels of i , $i \in [1, 5]$, will be the same.

$$L = L_1 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(7+5+3)^2}{(7^2+5^2+3^2)} = \frac{225}{83} = 2.71$$

$$L = L_2 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+4+3)^2}{(8^2+4^2+3^2)} = \frac{225}{89} = 2.53$$

$$L = L_3 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+2)^2}{(8^2+5^2+2^2)} = \frac{225}{93} = 2.42$$

$$L_4 = L_5 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+3)^2}{(8^2+5^2+3^2)} = \frac{256}{98} = 2.612$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{5}{\frac{1}{2.71} + \frac{1}{2.53} + \frac{1}{2.42} + \frac{1}{2.61} + \frac{1}{2.61}} = 2.57$$

7. Level of the $\sigma^2(p:d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, p:d) = 1$, $v \in [1, \sum g]$. For each unique i , there will be 1 count up to the total counts for that i .

$$L_1 = L_2 = L_3 = \frac{(\sum_{v=1}^{15} 1)^2}{\sum_{v=1}^{15} 1^2} = \frac{(1+1+1+1+1+1+1+1+1+1+1+1+1+1+1)^2}{(1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2)} = \frac{225}{15} = 15$$

$$L_4 = L_5 = \frac{(\sum_{v=1}^{16} 1)^2}{\sum_{v=1}^{16} 1^2} = \frac{(1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1)^2}{(1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2)} = \frac{256}{16} = 16$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{5}{\frac{1}{15} + \frac{1}{15} + \frac{1}{15} + \frac{1}{16} + \frac{1}{16}} = 15.38$$

8. Level of the $\sigma^2(i)$ term on facet of differentiation I. Since i is included in I, there is a single unique i for each I. $L = 1$

9. Level of the $\sigma^2(i \times d)$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, d) = |p_v|$, for $p = [8^*, 5^*, 3^*]$ with $v \in [1, 3]$; $g \in [1, 5]$ and *indicating missing data, similar to 6. Since i is not involved with the unbalanced nesting, all levels of i , $i \in [1, 5]$, will be the same.

$$L = L_1 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(7+5+3)^2}{(7^2+5^2+3^2)} = \frac{225}{83} = 2.71$$

$$L = L_2 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+4+3)^2}{(8^2+4^2+3^2)} = \frac{225}{89} = 2.53$$

$$L = L_3 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+2)^2}{(8^2+5^2+2^2)} = \frac{225}{93} = 2.42$$

$$L_4 = L_5 = \frac{(\sum_{v=1}^3 p_v)^2}{\sum_{v=1}^3 p_v^2} = \frac{(8+5+3)^2}{(8^2+5^2+3^2)} = \frac{256}{98} = 2.612$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{5}{\frac{1}{2.71} + \frac{1}{2.53} + \frac{1}{2.42} + \frac{1}{2.61} + \frac{1}{2.61}} = 2.57$$

10. Level of the $\sigma^2(i \times (p:d))$ term on facet of differentiation I. The sum of counts squared and sum of squared counts can be determined as follows:

$C(i, p:d) = 1$, $v \in [1, \sum g]$. Similar to 7, For each unique i , there will be 1 count up to the total counts for that i .

$$L_1 = L_2 = L_3 = \frac{(\sum_{v=1}^{15} 1)^2}{\sum_{v=1}^{15} 1^2} = \frac{(1+1+1+1+1+1+1+1+1+1+1+1+1+1+1)^2}{(1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2)} = \frac{225}{15} = 15$$

$$L_4 = L_5 = \frac{(\sum_{v=1}^{16} 1)^2}{\sum_{v=1}^{16} 1^2} = \frac{(1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1)^2}{(1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2)} = \frac{256}{16} = 16$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{5}{\frac{1}{15} + \frac{1}{15} + \frac{1}{15} + \frac{1}{16} + \frac{1}{16}} = 15.38$$

11. Level of the $\sigma^2(d)$ term on facet of differentiation D. Since d is included in D, there is a single unique D for each D.
 $L = 1$

12. Level of the $\sigma^2(p:d)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, p:d) = i_v$, where i_v is the number of items in v for $[g=A, v \in [1, 8]; g=B, v \in [1, 5]; g=C, v \in [1, 3]]$. Thus we have different levels of patients:doctors for each unique doctor.

$$L_A = \frac{(\sum_{v=1}^8 i_v)^2}{\sum_{v=1}^8 i_v^2} = \frac{(4+5+5+5+5+5+5+5)^2}{(4^2+5^2+5^2+5^2+5^2+5^2+5^2+5^2)} = 7.96$$

$$L_B = \frac{(\sum_{v=1}^5 i_v)^2}{\sum_{v=1}^5 i_v^2} = \frac{(4+5+5+5+5)^2}{(4^2+5^2+5^2+5^2+5^2)} = 4.97$$

$$L_C = \frac{(\sum_{v=1}^3 i_v)^2}{\sum_{v=1}^3 i_v^2} = \frac{(4+5+5)^2}{(4^2+5^2+5^2)} = 2.97$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{7.96} + \frac{1}{4.97} + \frac{1}{2.97}} = 4.52$$

13. Level of the $\sigma^2(i)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i) = |g_v|$, for $g=[8, 5, 3]$ and $v = 5$. There are g counts of each item $[1, 5]$. For example there are 7 counts of item $[1]$ for doctor A because 7 patients returned surveys including item 1 for doctor A.

$$L_A = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(7+8+8+8+8)^2}{(7^2+8^2+8^2+8^2+8^2)} = 4.99$$

$$L_B = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(5+4+5+5+5)^2}{(5^2+4^2+5^2+5^2+5^2)} = 4.97$$

$$L_C = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(3+3+2+3+3)^2}{(3^2+3^2+2^2+3^2+3^2)} = 4.9$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{4.99} + \frac{1}{4.97} + \frac{1}{4.90}} = 4.95$$

14. Level of the $\sigma^2(i \times d)$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i \times d) = |g_v|$, for $g=[8, 5, 3]$ and $v = 5$. There are g counts of each item $[1, 5]$. This is identical to the uncrossed items, $C(d, i)$.

$$L_A = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(7+8+8+8+8)^2}{(7^2+8^2+8^2+8^2+8^2)} = 4.99$$

$$L_B = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(5+4+5+5+5)^2}{(5^2+4^2+5^2+5^2+5^2)} = 4.97$$

$$L_C = \frac{(\sum_{v=1}^5 g_v)^2}{\sum_{v=1}^5 g_v^2} = \frac{(3+3+2+3+3)^2}{(3^2+3^2+2^2+3^2+3^2)} = 4.9$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{4.99} + \frac{1}{4.97} + \frac{1}{4.90}} = 4.95$$

15. Level of the $\sigma^2(i \times (p:d))$ term on facet of differentiation D. The sum of counts squared and sum of squared counts can be determined as follows:

$C(d, i \times (p:d)) = 1$, for $v=[\sum g]$ with $g=[A, B, C]$. For each doctor, there is a single unique count of items crossed with the nested patient:doctor. Thus to obtain the level for each unique facet of differentiation, g , we must sum over v , which is the sum of counts for each doctor (marginalizing over doctor).

$$L_A = \frac{(\sum_{v=1}^{39} 1)^2}{\sum_{v=1}^{39} 1^2} = \frac{1521}{39} = 39$$

$$L_B = \frac{(\sum_{v=1}^{24} 1)^2}{\sum_{v=1}^{24} 1^2} = \frac{576}{24} = 24$$

$$L_A = \frac{(\sum_{v=1}^{14} 1)^2}{\sum_{v=1}^{14} 1^2} = \frac{196}{14} = 14$$

$$L = \frac{|G|}{\sum_{g=1}^G \frac{1}{L_g}} = \frac{3}{\frac{1}{39} + \frac{1}{24} + \frac{1}{14}} = 21.62$$

Table 28. Denominators for each effect when the facet of differentiation changes when calculating G and D

| Variance Term | P:D Facet of Differentiation Denominator for G and D | I Facet of Differentiation Denominator for G and D | D Facet of Differentiation Denominator for G and D |
|----------------------------|---|---|---|
| $\sigma^2(d)$ | 1 | 2.573 | 1 |
| $\sigma^2(p:d)$ | 1 | 15.385 | 4.520 |
| $\sigma^2(i)$ | 4.776 | 1 | 4.951 |
| $\sigma^2(i \times d)$ | 4.776 | 2.573 | 4.951 |
| $\sigma^2(i \times (p:d))$ | 4.776 | 15.385 | 21.624 |

These values can now be used to calculate the G and D coefficients. We will complete the one for P:D (Table 29). Note that the values for G and D are slightly lower (.700 and .675, respectively) than those with no missing data (.752 and .718, respectively), which is to be expected.

Table 29. Calculating G and D Coefficients for P:D Narayanan et al. (2010) missing data

| Generalizability Coefficient | Facet of Differentiation = P:D |
|------------------------------|---|
| Generalizability (G) | $G = \text{Tau} / (\text{Tau} + \text{EV}(G))$ $\text{Tau} = \text{VC}(P:D) + \text{VC}(D)$ $\text{Tau} = .083 + .001 = .084$ $\text{EV}(G) = \text{VC}(P:I:D) / \text{Level}_{P:D}(P:I:D)$ $= .170 / 4.776$ $= .036$ $G = .084 / (.084 + .036)$ G = .700 |
| Dependability (D) | $D = \text{Tau} / (\text{Tau} + \text{EV}(D))$ $\text{EV}(D) = (.021 / 4.776) + (0 / 4.776) + (.170 / 4.776)$ $= .040$ $D = .083 / (.083 + .040)$ D = .675 |

Limitations of this Procedure

1. All effects are assumed to be random. This is a very difficult issue to get around when data are unbalanced/have missing data points.
2. Sample sizes are uncorrelated with effects.
3. All effects (except the grand mean) are uncorrelated with each other, have means of 0 and have variances.

While the only design that has been gone through in great detail in this tutorial is the $i \times (p:d)$ design. ALL other designs (crossed, nested, etc. – perhaps using many more terms) are theoretically sound with tests passing for many different synthetic datasets (see <https://github.com/tylerjsmith111/GeneralizIT>). The only limiting factor for the automated calculation of generalizability coefficients for any study design is the data analyst's ability to properly specify the linear random effects relationships present in a given study (for example in the $i \times (p:d)$ -> $X = \mu + \sigma_i^2 + \sigma_d^2 + \sigma_{p:d}^2 + \sigma_{id}^2 + \sigma_{pi:d}^2$

Conclusion

The utility in creating a generalizability program allows for: 1) missing data, 2) unbalanced data, 3) providing both generalizability and dependability coefficients, and 4) D-values for designs with no missing data is very useful addition to the toolbox of researchers. There is currently not a freely-available one that does all of these things.

References

Bloch, R. & Norman, G. (2023). *G String V User Manual*. Hamilton, Ontario, Canada.

Bloch, R. & Norman, G. (2012). Generalizability theory for the perplexed: A practical introduction and guide: AMEE Guide No. 68. *Medical Teacher*, 34 (11), 960-992. DOI: 10.3109/0142159X.2012.703791

Brennan, R. L. (2001a). *Generalizability Theory*. New York: Springer.

B Brennan, R. L. (2001b). *Manual for urGENOVA* (Version 2.1) (Iowa Testing Programs Occasional Paper Number 49). Iowa City, IA: Iowa Testing Programs, University of Iowa.

Riesch, A.M., Swaminathan, H., Welsh, M. & Chafouleas, S.M. (2014). Generalizability theory: A practical guide to study design, implementation, and interpretation, *Journal of School Psychology*, 52 (1), 13-35.

Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics*, 9(2), 226-252.

Narayanan, A., Greco, M., & Campbell, J.L. (2010). Generalisability in unbalanced, uncrossed and fully nested studies. *Medical Education*, 44(4), 367-387.

Water Quality Division (2010). *Procedures to Implement the Texas Surface-Water-Quality Standards: TCEQ RG-194*, p. 81.

Appendix A: Harmonic Mean Discussion

Harmonic means are an average. It is calculated by (i) taking the sum of the reciprocals of each value in a data series, (ii) dividing the sum by the number of values in the data series, (that is the averaging part) and (iii) taking the reciprocal of that number. Mathematically equivalently: taking the sum of the number of values in the data series and dividing it by the sum of the reciprocals of each value in a data series.

Harmonic means are most useful when you are dealing with a set of values based on rates and ratios (i.e., working with reciprocal relationships). Ratios (such as speed (km/hr); Price/Earnings ratios; and Nested variables #items/form, #patients/doctor) are such instances. Harmonic means are thus useful in unbalanced data designs where there are non-equivalent numbers of items nested within another factor (e.g., i:h or p:h). Clearly, these are ratios.

These websites are useful to differentiate between arithmetic, geometric and harmonic means.

<https://www.datacamp.com/tutorial/harmonic-mean>

<https://ryxcommar.com/2023/01/13/intuitive-explanation-of-arithmetic-geometric-harmonic-mean/>

Using the simple example of: What is the average speed of a vehicle that drives out 60 km at 60/km/hr and back the same distance at 20 km/hr?

The arithmetic mean would be $(60+20)/2\text{hrs} = 80/2 = 40$ km/hr. However, this is NOT the average speed travelled over a fixed distance (i.e., “there and back”). The problem is that on the way back, the driver would have only covered only 1/3 of the

distance back since they are going at 20/km/hr. To cover the while distance, they would have to drive 3 hours. We CAN mentally work this through and take the weighted arithmetic mean to get at the correct average speed:
 $(60+20+20+20)/4 \text{ hrs} = 120/4 = 30 \text{ km/hr}$ is the average speed for the fixed distance when travelling 60 km/hr out and 20 km/hr back.

While this worked with a simple example, a much more elegant way to do this with many numbers that don't evenly work out for weighting, is to take the reciprocal of the values ($1/60$ and $1/20$), then average them $(0.016667 + 0.05)/2 = 0.033333$. Then to "get back" to our original units of km/hr, we take the reciprocal of this value $(1/0.033333) = 29.99999 = 30$. We could also skip a step and when we get to $0.066667/2$, simply take the reciprocal of it: $2/0.0666667 = 29.999999 = 30$. In looking at the reciprocal values, you can see that the larger the value to start with, the smaller its reciprocal value. By taking the average of the reciprocals, you are equally weighting each value before averaging them. This dampens the effect of larger values in the data set that is exerted with the arithmetic mean.

For example: The arithmetic mean of Brennan's (2001a) data set (Table 7.5, p. 224) with 3 h's (let's refer to them as Forms) and within each Form there are 2, 4, and 2 items. So, in effect we have three ratio values: 2items/1 form, 4items/1 form, 2items/1 form.

The arithmetic mean of these values is: $(2+4+2)/3 = 8/3 = 2.667$

The harmonic mean of these values is: $[3/(\frac{1}{2} + \frac{1}{4} + \frac{1}{2})] = (3/(0.5+0.25+0.5)) = 3/1.25 = 2.4$

Why would we want the arithmetic versus the harmonic means?

The arithmetic mean (2.667) tells us that there are, on average, 2.667 items for each form for this data set. There are 8 persons (p) who take all the items in all the forms. If we take 2.667 and multiply it by the number of persons (8) we get 21.336 "counts" estimated for each form. If we multiply this value by the number of forms (3), we get 64 – the number of counts in the entire data set FOR PERSONS. In addition, for each person, (person 1 has 2.667 items for Form 1, 2.667 items for Form 2 and 2.667 items for Form 3) returns 8 counts ($2.667 * 3$) per person. This symmetry is useful to know and works when there are no missing data.

The harmonic mean "weights" the forms equally so that the effect of forms "assumes" they all have the same number of items in them (are pseudo-balanced). For example, Form 2 has 4 items in it; if we weight Forms 1 and 3 by doubling their number of items, then we can calculate the arithmetic average of:

Form h1(2+2)+

Form h2(4)+

Form h3(2+2)

=12

5 values in the data set

$12/5 = 2.4$. This equally weights the contribution of each of the forms when using the term in the denominators of the Generalizability equations.