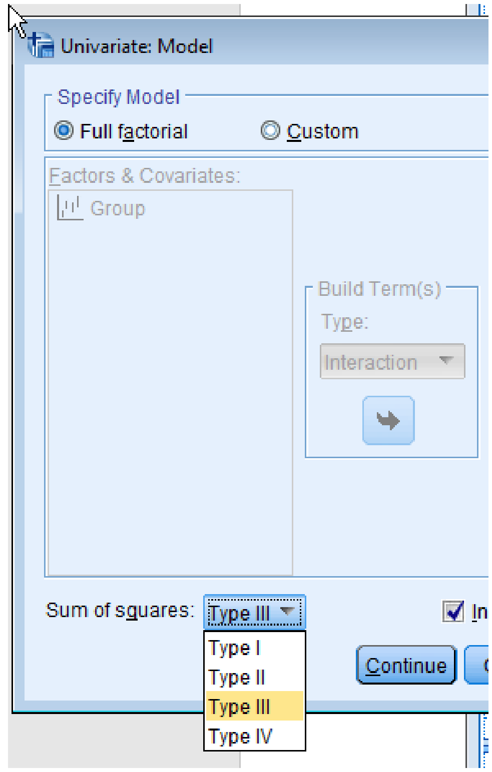
Memo

To: Colleagues
From: Roger Bakeman
Date: August 22, 2020
Re: What Type Sums of Squares Should You Use: I, II, or III?



The devil is in the defaults. When confronted with options in statistical computer packages like SPSS, and uncertain what different options do, we are likely to trustingly assume that the designers knew what they were doing and so, a bit mindlessly, accept the default.

One such default is the Sum of Squares option in SPSS's **General Linear Model** procedure. The default is Type III, but is that always the best choice? Are there times when you should specify Type I or Type II instead? (I won't discuss Type IV, which is a variation of Type III developed for designs with missing cells.)

Does It Matter?

Under specific circumstances, it doesn't matter which you select. When your model:

- is balanced (i.e., all cells contain the same number of scores),
- includes only one between-subjects factor (even if cell *n*s are unequal),
- includes only one between- and one or more within-subjects factor or factors, or
- consists only of within-subjects factors (cell *n*s are equal by definition),

then Type I, II, and III sums of squares will be identical.

When Might It Matter?

However, if your model contains two or more between-subjects factors and cell *n*s are not equal, Type I, II, and III sums of squares give different results. Different communities of users and different statistical traditions have different preferences. Analysis of variance traditionalists, experimentalists for whom all variables are manipulated and controlled, and SPSS and SAS users generally, tend to prefer Type III. Multiple regression traditionalists, researchers whose variable are often natural and whose studies are correlational, and R users generally tend to prefer Type I (although R users increasingly are leaning toward Type II; Mangiafico, 2015).

Remember, if sums of squares are different, the statistics that depend on them—partial and generalized eta squareds, *p* values, *F* ratios, and mean squares– will be as well.

Differences among the types are easy to understand if you keep in mind that the general linear model consists of terms. For example, the terms for a two-way, AB model are A main effect, B main effect, A×B interaction, and error (s/AB, subjects within AB). For details and other examples see Bakeman (1992) and Bakeman and Robinson (2005).

Here is the difference:

Type I sum of squares is **sequential.** Terms are added to the model in the order specified. The sum of squares for each is the increase in the sum of squares—that is, the reduction in the error sum of squares—when that term is added. The total sum of squares—the grand mean subtracted from each score, squared, and summed—is partitioned exactly, just as Cohen (1968) described in his ground-breaking article, *Multiple Regression as a General Data-Analytic System*. That is, the sums of squares for the terms add up to the total sum of squares.

Type II and Type III sum of squares are **partial.** For Type III, the sum of squares for each term is the increase in the sum of squares when the term under consideration is added last, after all other terms have been entered. Thus, it takes into account (controls for) all other terms in the model (i.e., all main effects and interactions).

For Type II, the sum of squares for each term is the increase in the sum of squares when the term under consideration is added to a model that excludes all higher-order terms containing that term. For example, for an ABC model, A would be added to a model containing B, C, and BC, while excluding AB, AC, and ABC. Thus, it takes into account (controls for) all other terms in the model of which it is not a part.

Consequently—and unlike Type I—Type II and III sums of squares do not sum exactly to the total sum of squares, a circumstance that has both pros and cons (Reid, 2009).

The SPSS algorithm manual describes the computations exactly, expressed mainly with matrix algebra. To give you a flavor, here is a quote: "The Type III sum of squares for an effect F can best be described as the sum of squares for F adjusted for effects that do not contain it, and orthogonal to effects (if any) that contain it" (IBM SPSS Statistics 24 Algorithms, 2016, p. 1061).

When to Use Type I

With Type I sum of squares, order matters. When cell *n*s are unequal, specifying variables in different orders gives different results. Thus, Design = A B A*B and Design = B A A*B are not the same. This makes Type III the "safe" choice; you don't need to commit to an order. But if you want to test one effect controlling for another, then you should use Type I. For example, imagine you have a sample of boys and girls who have had three types of childcare. You are primarily interested in whether type of childcare affects performance on a social skills test and you know that there tend to be sex differences in how children perform on this test. Thus, you specify gender first, to control for it, and then add type of childcare.

In sum, if your model includes more than one between-subjects variable, and you have reason to want to consider the variables in a particular order, controlling for the effects of the variable listed first, you should consider Type I. It wouldn't hurt to run analyses with a different order, or with Type II and Type III specified, just for comparison; often the differences will be slight. Note, I am not recommending that you try different specifications and pick the one that you like best, just that it can be informative to note differences, if any—after you have decided a priori which type sum of squares to use.

When to Use Type II

As noted earlier, Type III is the default for SPSS, SAS, and several other statistical packages, and is the traditional choice. Recently Langsrud (2003) had recommended that Type II and not Type III be used when cells sizes are unequal (as noted earlier, when cell sizes are equal it doesn't matter). I find his arguments, although technical, well-reasoned and compelling—and, as noted earlier, increasingly R users seem to be agreeing. Despite my previous multiple-regression-based bias against Type III and for Type I, and absent an explicit rationale for Type I given the variables employed, I would now say Type II is probably the better choice. You can always compare with Type III and note any differences; often I suspect they will be slight.

Inspiration. This memo was inspired by my belief that knowing is better than not.
As the Notorious B.I.G. said, "If you don't know, now you know."

Acknowledgement. Thanks to Daryl Nenstiel and Frank Haist for comments on an earlier draft.

An Example

For this example, the data—social skills scores—are analyzed with a 2 × 3, sex by childcare type between-subjects model. This is not an actual data set but one that I jiggered a bit for illustrative purposes. Here are the number of scores in each cell and their means.

N	Type A	Type B	Type C	Total
Boys	5	5	9	19
Girls	7	9	9	25
Total	12	14	18	44

M	Type A	Type B	Type C	Total
Boys	35.0	40.6	43.9	40.7
Girls	40.1	47.6	57.8	49.2
Total	38.0	45.1	50.8	45.5

On the next page are the SPSS results, but here are the highlights:

	Term	Type I, sex then type	Type I, type then sex	Type II	Type III
SS	Sex	776	938	938	771
	Type	1,352	1,190	1,352	1,266
	S x T	163	163	163	163
η_p^2	Sex	.094	.111	.111	.093
	Type	.153	.137	.153	.144
	S x T	.021	.021	.021	.021
p	Sex	.055	.036	.036	.056
	Type	.043	.061	.043	.052
	S x T	.665	.665	.665	.665

Sums of squares, partial eta squareds, and *p* values all vary a bit, depending on the type. With Type III, neither the sex nor the type main effect achieves statistical significance, $\alpha = .05$. With Type II, both are statistically significant, $p < .05$. And with Type I the one with type entered first is statistically significant, $p < .05$, but the one with type entered after sex is not.

Nonetheless, no matter which side of the .05 "cliff" the *p* values fall, the effects sizes are all relatively similar. For the sex main effect, all fall in the medium range per Cohen's benchmarks. But for the type main effect, whether it is medium or just barely large depends on the type.

Here are analysis of variance results with Type I, II, and III sum of squares (SS) specified, with sex-then-type and type-then-sex orders for the two analyses with Type I sum of squares.

Tests of Between-Subjects Effects

Source	Type I SS: sex, type	df	MS	F	p	Partial η2
Corrected Model	2,290	5	458	2.32	.062	.234
Intercept	91,091	1	91091	461.24	.000	.924
Sex	776	1	776	3.93	.055	.094
Type	1,352	2	676	3.42	.043	.153
Sex * Type	163	2	81	0.41	.665	.021
Error	7,505	38	197			
Total	100,886	44				
Corrected Total	9,795	43				

Source	Type I SS: type, sex	df	MS	F	p	Partial η2
Corrected Model	2,290	5	458	2.32	.062	.234
Intercept	91,091	1	91091	461.24	.000	.924
Type	1,190	2	595	3.01	.061	.137
Sex	938	1	938	4.75	.036	.111
Sex * Type	163	2	81	0.41	.665	.021
Error	7,505	38	197			
Total	100,886	44				
Corrected Total	9,795	43				

Source	Type III SS	df	MS	F	p	Partial η2
Corrected Model	2,290	5	458	2.32	.062	.234
Intercept	80,127	1	80127	405.72	.000	.914
Sex	771	1	771	3.90	.056	.093
Type	1,266	2	633	3.21	.052	.144
Sex * Type	163	2	81	0.41	.665	.021
Error	7,505	38	197			
Total	100,886	44				
Corrected Total	9,795	43				

Source	Type II SS	df	MS	F	p	Partial η2
Corrected Model	2,290	5	458	2.32	.062	.234
Intercept	91,091	1	91091	461.24	.000	.924
Sex	938	1	938	4.75	.036	.111
Type	1,352	2	676	3.42	.043	.153
Sex * Type	163	2	81	2.32	.665	.021
Error	7,505	38	197			
Total	100,886	44				
Corrected Total	9,795	43				

The total SS for the dependent variable—called the Corrected Total in the SPSS output—is the same for all four analyses, but the Sex, Type, Sex × Type, and error term sum of squares (SS) sum to the corrected total SS only with Type I sum of squares. (You can ignore the Corrected Model, Intercept, and Total SS, but in the SPSS Model dialog box leave *Include Intercept in Model* checked.)

References

Bakeman, R. (1992). *Understanding social science statistics: A spreadsheet approach.* Hillsdale, NJ: Erlbaum.

Bakeman, R., & Robinson, B. F. (2005). *Understanding statistics in the behavioral sciences*. Mahwah, NJ: Erlbaum.

Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin, 70*, 426–443.

IBM SPSS Statistics 24 Algorithms. (2016). PDF retrieved from ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/24.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf.

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II Instead of Type III sums of squares, *Statistics and Computing, 13*, 163–167. https://link.springer.com/content/pdf/10.1023/A:1023260610025.pdf

Mangiafico, S. S. (2015). *An R Companion for the Handbook of Biological Statistics*, version 1.3.2. rcompanion.org/rcompanion/. (Pdf version: rcompanion.org/documents/RCompanionBioStatistics.pdf.)

Reid, N. (2009). Types of sums of squares. University of Toronto. http://www.utstat.utoronto.ca/reid/sta442f/2009/typeSS.pdf (Initially Analysis of Functional Neuro Images, National Institute of Mental Health, http://afni.nimh.nih.gov/sscc/gangc/SS.html, now an inactive page.)