
Dan J. Putka & Rodney A. McCloy
Human Resources Research Organization

This document explains how to estimate variance components in SPSS and SAS for a variety of measurement designs that involve ratings. Variance components serve as the building blocks of reliability coefficients discussed in the literature on generalizability theory and intraclass correlations (Cronbach, Gleser, Nanda, & Rajaratnam, 1972; DeShon, 2002; McGraw & Wong, 1996).

Setting Up Your Data

Whether you use SPSS or SAS, your data need to be formatted a specific way (univariate format) to generate variance component estimates. What distinguishes a univariate format from a multivariate format is that the primary outcome variable of interest (i.e., ratings) appears in only one column. Univariate formats will result in data sets with multiple records (rows) per object of measurement (e.g., persons), whereas multivariate formats will typically have only one record per object of measurement. Given that organizational researchers are more likely accustomed to dealing with data in multivariate formats, Tables 1 through 4 show examples of multivariate and univariate formats of data sets produced by four types of measurement designs.

Table 1. Example of Multivariate and Univariate Formats for Data Arising from a Person x Rater (p x r) Measurement Design

<table>
<thead>
<tr>
<th>Multivariate Format</th>
<th>Univariate Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id rater1 rater2</td>
<td>person_id rater_id rating</td>
</tr>
<tr>
<td>1 3 4</td>
<td>1 1 3</td>
</tr>
<tr>
<td>2 4 2</td>
<td>1 2 4</td>
</tr>
<tr>
<td>3 5 1</td>
<td>2 1 3</td>
</tr>
<tr>
<td></td>
<td>2 2 2</td>
</tr>
<tr>
<td></td>
<td>3 1 5</td>
</tr>
<tr>
<td></td>
<td>3 2 1</td>
</tr>
</tbody>
</table>

1 This document can be found online at: [http://www.humrro.org/djp_archive/Variance_Component_Estimation_in_SPSS_and_SAS.pdf](http://www.humrro.org/djp_archive/Variance_Component_Estimation_in_SPSS_and_SAS.pdf) and was first posted in February of 2008. An earlier version of this document was presented in McCloy, R. A., & Putka, D. J. (2004, April). Estimating interrater reliability: Conquering the messiness of real-world data. Master Tutorial conducted at the 19th Annual Society for Industrial and Organizational Psychology Conference, Chicago, IL.
Table 2. Example of Multivariate and Univariate Formats for Data Arising from Rater:Person (r:p) Measurement Design

<table>
<thead>
<tr>
<th>Multivariate Format</th>
<th>Univariate Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id</td>
<td>rater1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Example of Multivariate and Univariate Formats for Data Arising from an Ill-Structured Design with Two Raters Per Person

<table>
<thead>
<tr>
<th>Multivariate Format</th>
<th>Univariate Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id</td>
<td>rater1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Note. Such designs frequently arise in the study of performance ratings, interview ratings, and assessment center ratings in applied organizational research and practice (see Putka, Le, McCloy, & Diaz, in press).

Table 4. Example of Multivariate and Univariate Formats for Data Arising from a Person x Rater x Item (p x r x i) Measurement Design

<table>
<thead>
<tr>
<th>Multivariate Format</th>
<th>Univariate Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id</td>
<td>rater1_item1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fortunately, it is a relatively simple matter to convert a data set from multivariate format to univariate format. Within SPSS, one can use the VARTOCASES command to restructure the multivariate data file in Tables 1 through 4 into univariate format.

**Table 1: Multivariate-to-Univariate Conversion**

```
VARSTOCASES
/MAKE rating FROM rater1 rater2
/INDEX = rater_id(2)
/KEEP =  person_id
/NULL = DROP.
```

SPSS will make a new “rating” variable that reflects values for the current rater1 and rater2 variables.

SPSS will make a new “rater_id” variable that takes on two values (1, 2) corresponding to the rater who provided the given value on the new rating variable.

SPSS will keep the person_id variable to indicate which values in the new rating variable are for each person, and drop any person_id-x-rater_id combinations that are missing values on the rating variable.

**Table 2 and 3: Multivariate-to-Univariate Conversion**

```
VARSTOCASES
/MAKE rating FROM rater1 rater2 rater3 rater4
/INDEX = rater_id(4)
/KEEP =  person_id
/NULL = DROP.
```

**Table 4: Multivariate-to-Univariate Conversion**

```
VARSTOCASES
/MAKE rating FROM rater1_item1 rater1_item2 rater2_item1 rater2_item2
/INDEX = rater_id(2) item_id(2)
/KEEP =  person_id
/NULL = DROP.
```

These examples could easily be expanded to designs that include more raters or items. We use a small number of raters and items here for illustration purposes only.

*Estimating Variance Components for Single-Facet Rating Designs*

Next, we present SAS and SPSS syntax for estimating variance components that arise from a variety of single-facet rating designs. In these designs, persons are the objects of measurement, and rater_id is treated as a random facet of measurement.
1. p x r design (also applicable for the ill-structured design)

*SAS-------------------------

PROC MIXED method=REML;
class person_id rater_id;
model rating =;
random person_id rater_id;
run;

Indicate the facets of your measurement design and your object of measurement on the CLASS statement.

Indicate your primary outcome variable (e.g., rating) before the "equals" sign. If fixed effects were part of your measurement design, they would be listed after the "equals" sign and resulting variance component estimates would be conditional on the fixed effects.

Indicate all variance components that can be estimated based on your measurement design (except the highest order interaction—see note below). This code will produce estimates for Var(person), Var(rater), and Var(residual).

*SPSS----------------------

VARCOMP
rating BY person_id rater_id
/RANDOM = person_id rater_id
/METHOD = REML
/DESIGN = person_id rater_id
/INTERCEPT = INCLUDE .

In SPSS, the first line after the VARCOMP command should indicate your primary outcome variable (e.g., rating) BY your object of measurement and facets of your measurement design.

As in SAS, SPSS provides you with a variety of methods for estimating variance components.

The /DESIGN statement in SPSS VARCOMP is similar to the RANDOM statement in SAS PROC MIXED. Again, indicate all variance components that can be estimated based on your measurement design (except the highest order interaction—see note below).

Note. You do not have to indicate the highest-order interaction term in your measurement design on the RANDOM (SAS) or /DESIGN (SPSS) statements, because it is confounded with residual error (which is always estimated by these models). Your estimate for the variance component for the highest-order interaction (as well as residual error) is the variance component estimate for residual error produced by these programs.
2. r:p design

*SAS-------------------------
PROC MIXED method=REML;
class person_id rater_id;
model rating =;
random person_id;
run;

*SPSS-------------------------
VARCOMP
rating  BY person_id rater_id
/RANDOM = person_id rater_id
/METHOD = REML
/DESIGN = person_id
/INTERCEPT = INCLUDE .

3. p:r design

*SAS-------------------------
PROC MIXED method=REML;
class person_id rater_id;
model rating =;
random rater_id;
run;

*SPSS-------------------------
VARCOMP
rating  BY person_id rater_id
/RANDOM = person_id rater_id
/METHOD = REML
/DESIGN = rater_id
/INTERCEPT = INCLUDE .

Estimating Variance Components for Two-Facet Rating Designs

We now present SAS and SPSS syntax for estimating variance components that arise from a variety of two-facet rating designs (raters and items). In these designs, persons are the objects of measurement, and rater_id and item_id are treated as random facets of measurement.

1. p x r x i design

*SAS-------------------------
PROC MIXED method=REML;
class person_id rater_id item_id;
model rating =;
random person_id rater_id item_id
  person_id*item_id person_id*rater_id
  rater_id*item_id;
run;

*SPSS-------------------------
VARCOMP
rating  BY person_id rater_id item_id
/RANDOM = person_id rater_id item_id
/METHOD = REML
/DESIGN = person_id rater_id item_id
person_id*item_id person_id*rater_id
rater_id*item_id
/INTERCEPT = INCLUDE .

This code will produce estimates for Var(person), Var(rater), Var(item), Var(person-x-item), Var(person-x-rater), Var (rater-x-item), and Var(residual).
2. \((r:p) \times i\) design

*SAS-------------------------

PROC MIXED method=REML;
class person_id rater_id item_id;
model rating =;
random person_id rater_id(person_id)
item_id person_id*item_id;
run;

*SPSS-------------------------

VARCOMP
rating BY person_id rater_id item_id
/RANDOM = person_id rater_id item_id
/METHOD = REML
/DESIGN = person_id rater_id(person_id) item_id
person_id*item_id
/INTERCEPT = INCLUDE .

3. \((p:r) \times i\) design

*SAS-------------------------

PROC MIXED method=REML;
class person_id rater_id item_id;
model rating =;
random rater_id person_id(rater_id)
item_id rater_id*item_id;
run;

*SPSS-------------------------

VARCOMP
rating BY person_id rater_id item_id
/RANDOM = person_id rater_id item_id
/METHOD = REML
/DESIGN = rater_id person_id(rater_id) item_id
rater_id*item_id
/INTERCEPT = INCLUDE .

In SAS and SPSS, nesting is indicated with parentheses. Thus, if factor A was nested within factor B, we would use A(B).

References


