Tutorial for the Use of *ConfirmatoryANOVA.exe*

Rebecca M. Kuiper

Utrecht University *

This tutorial will describe and illustrate the options available in *ConfirmatoryANOVA.exe*. This program is free of use. However, when results obtained with this program are published, please refer to Kuiper and Hoijtink (2010) and/or Kuiper, Klugkist, and Hoijtink (2010). In the program, the following methods can be performed:

- $\bar{F}$ test
- Order-Restricted Information Criterion (ORIC)
- Bayesian Model Selection (BMS)

The model used in this tutorial and in the software is the ANOVA model:

$$y_{ij} = \mu_i + \epsilon_{ij}, \tag{1}$$

where y_{ij} is the jth observation ($j = 1, \ldots, n_i$) of the dependent variable for group i ($i = 1, \ldots, k$), μ_i is the mean of group i, and ϵ_{ij} is the error term. The error terms are independently and normally distributed, with expected value 0 and variance σ^2, that is, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$.

*Rebecca M. Kuiper, Department of Methodology and Statistics, Utrecht University. Correspondence concerning this article should be addressed to Rebecca M. Kuiper, Department of Methodology and Statistics of the Utrecht University, P.O.Box 80140, 3508TC Utrecht, The Netherlands. E-mail: `R.M.Kuiper@uu.nl`. This research is financed by the Dutch organization for scientific research NWO-VICI-453-05-002.

Modification Input Files

No matter what analysis should be performed, two text files have to be modified (such that they apply to your data), namely *Input.txt* and *Data.txt*.

It should be noted that:

- The names of the text files are fixed and cannot be changed. These files have to be text files (also known as ASCII files).
- The format of these files should not be changed, that is, do not add empty lines and do not delete lines containing labels.
- The data in *Data.txt* should be complete, that is, missing data are not allowed. Furthermore, a 'dot (".") is used as decimal separator, not a comma (",").

First Half of *Input.txt*

First of all, you must denote which analyses should be performed. This has to be done in *Input.txt*. A certain analysis will be performed if in the line below the name of that analysis a 1 is filled in. It will not be performed, when a 0 is filled in.

When ORIC and the $\bar{F}$ test should be performed and BMS should not, the first half of *Input.txt* should look as follows (when using the default values for the seed value and number of iterations):

```
Seed value and number of iterations (> 0) for Fbar test, ORIC, and BMS
123 100000 100000 500000
Perform F bar test, ORIC, BMS (1 = yes, 0 = no)
1 1 0
```

We will come back to the seed value and number of iterations later on the tutorial in the section called "Set the Seed Value and Number of Iterations".

Data.txt

Second of all, group membership and the corresponding data must be given in *Data.txt*, where in the first column the group numbers must be given and in the second column the corresponding data point y_{ij}. The order of the group numbers and data points does not matter as long as the group number corresponds to the data point in the same row. The following are three examples of how *Data.txt* could look:

1	3.58	1	3.58	3	1.39
1	-0.15	2	1.67	2	1.85
...	...	3	1.39	5	4.58
2	1.67	4	1.57	5	1.38
2	1.85	5	1.38	3	4.53
...	...	1	-0.15	4	2.97
3	1.39	2	1.85	1	-0.15
3	4.53	3	4.53	2	1.67
...	...	4	2.97	4	1.57
4	1.57	5	4.58	1	3.58
4	2.97	...	...	...	...
...	...	...	...	...	...
5	1.38	...	...	...	...
5	4.58	...	...	...	...
...	...	...	...	...	...

Note that a dot (".") should be used as decimal separator. When a comma (",") is used, only the number proceeding it is read (e.g., "3,58" is read as "3").

The group numbers do not need to be sequential. For example, if you have a SPSS or Excel file with data for 10 groups and in the current analysis you want to compare only 5 groups (which are not the first five). In that case, you can just copy the appropriate data from the SPSS or Excel file to a text file without adjusting the group numbers. In the software, the group numbers will be made sequential; moreover, the order of the groups remain the same. For example, if you have data with group numbers 1, 4, 5, 6, and 8 (whether the data are in order or not), these will become group numbers 1, 2, 3, 4, and 5, respectively. Note that, in specifying the restrictions (see the next sections), you need to use the adjusted sequential group numbers. In addition, the output will be given in terms of these adjusted group numbers.

From the data the number of groups and the number of observations per group are determined.

Basic Elements of Writing Constraints

In performing an $\bar{F}$ test, in determining ORIC or in doing BMS, all the hypotheses of interest must be given explicitly. So, one must specify (order-restricted) hypotheses, that is, hypotheses with restrictions of the form $\mu_i - \mu_{i'} \geq 0$ for some $i, i' = 1, \ldots, k$.

Note that it is also possible to specify a set of hypotheses without the classical null $H_0 : \mu_1 = \ldots = \mu_k$ and/or alternative $H_A : \mu_1, \ldots, \mu_k$. However, we recommend to include the alternative H_A (when doing model selection), since it can be used to protect against choosing a weak hypothesis (Kuiper and Hoijtink, 2010). Note that one should include H_0 only when there is real interest in H_0.

In the $\bar{F}$ test, the order-restricted hypotheses (like $H_1 : \mu_1 > \mu_2 > \mu_3$) are tested against H_0 ("ordered alternative") and H_A ("ordered null"). If the classical null and/or the alternative are included in the set of hypotheses, the classical null will be tested against the classical alternative.

When using the ORIC or confirmatory BMS, several models/hypotheses are compared to each other, for example, $H_1 : \mu_1 > \mu_2 > \mu_3$, $H_2 : \mu_1 > \mu_2 = \mu_3$ and $H_3 : \mu_1,\ \mu_2,\ \mu_3$.

The basic elements for writing down the hypotheses of interest are:

1. Representation of an equality sign (=)

   Suppose the hypothesis of interest is $\mu_5 = \mu_3$, that is, $\mu_5 = \mu_3,\ \mu_1,\ \mu_2,\ \mu_4$. The ordering of the group numbers in this restriction is represented by: 5 3 1 2 4. The restriction is represented by: 1 1 0 0 0, where the 1s indicate that mean 5 and 3 belong to set 1 and are equal to each other, and where 0 indicates that the corresponding mean is unrestricted. N.B. in a restriction the first set of means is always labeled as 1, the second as 2 (and so on).

2. Representation of a greater than sign (>)

   Suppose the hypothesis of interest is $\mu_1 > \mu_3,\ \mu_2,\ \mu_4,\ \mu_5$. The ordering of the group numbers in this restriction is represented by: 1 3 2 4 5. The restriction is represented by: 1 -3 0 0 0, where -3 means that mean 3 is smaller than mean 1. Thus, it represents $\mu_3 < \mu_1$, which is equal to $\mu_1 > \mu_3$. Here again 1 indicates that mean 1 belongs to set 1. Because of the inequality restriction between mean 1 and 3, mean 3 belongs to set 2 (the importance of this will be made clear in the next section).

3. Representation of a smaller than sign (<)

   Suppose the hypothesis of interest is $\mu_1 < \mu_2,\ \mu_3,\ \mu_4,\ \mu_5$. The ordering of the group numbers in this restriction is represented by: 1 2 3 4 5. The restriction is represented by: 1 -1 0 0 0, where -1 means that mean 2 is greater than mean 1. Thus, it represents $\mu_2 > \mu_1$, which is equal to $\mu_1 < \mu_2$.

4. Representation of a free parameter

   When a mean is not constrained at all, it is called a free parameter. A free parameter is represented by a '0'.
   Note that, in a hypothesis, a free parameter is represented by ", μ_i,", or when it is the last element in the restriction by ", μ_i".

Every hypothesis can be represented by these basic elements in one or more restrictions. For example, $\mu_5 > \mu_3 < \mu_1,\ \mu_2 < \mu_4$, can be represented by the restrictions:
$\mu_5 > \mu_3,\ \mu_1,\ \mu_2,\ \mu_4,$
$\mu_5,\ \mu_3 < \mu_1,\ \mu_2,\ \mu_4,$
$\mu_2 < \mu_4,\ \mu_5,\ \mu_3,\ \mu_1,$
which can be represented by:

Ordering of means in restriction
5 3 1 2 4
5 3 1 2 4
2 4 5 3 1
(Order) Restrictions
1 -3 0 0 0
0 1 -1 0 0
1 -1 0 0 0

So, the ordering of the means in a certain restriction always consists of the numbers 1 to 'the number of groups' (in the example, 5) and each number is used only once.

Combinations of Basic Elements

Often the hypothesis of interest can be represented in a smaller number of restrictions than when using only the basic elements. The following shortcuts can be used:

1. $\mu_5 = \mu_3 = \mu_1,\ \mu_2 = \mu_4$

   Because of the equality constraints ("="), means 5, 3 and 1 belong to set 1. Therefore, means 2 and 4 belong to set 2. Thus, this hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 1 2 2

2. $\{\mu_5 = \mu_3 = \mu_1\} > \{\mu_2 = \mu_4\}$

   Because of the equality constraints, mean 5, 3 and 1 belong to set 1, and means 2 and 4 belong to set 2. Because of the constraint between mean 1 and 2, mean 2 belongs implicitly to set 2. This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 1 -3 2

   The 1s indicate the equality constraints between mean 5, 3 and 1, the -3 represents the inequality constraint ">" between mean 1 and 2, and the 2 indicates the equality constraints between mean 2 and 4 (because mean 2 implicitly belongs to set 2). N.B. the restrictions $\mu_5 = \mu_1,\ \mu_5 > \mu_2,\ \mu_3 > \mu_2,\ \mu_5 > \mu_4,\ \mu_3 > \mu_4,\ \mu_1 > \mu_4$ do not have to stated explicitly, these will hold since it holds that $\mu_5 = \mu_3,\ \mu_3 = \mu_1,\ \mu_1 > \mu_2$ and $\mu_2 = \mu_4$.

3. $\mu_5 = \mu_3 > \mu_1 > \mu_2 = \mu_4$

   This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 -3 3

The 1s indicate the equality constraint between mean 5 and 3. The -3s represents the inequality constraint ">" between mean 3 and 1 and between 1 and 2. Note that mean 1 implicitly belongs to set 2 and mean 2 implicitly to set 3. Therefore, the equality constraint between mean 2 and 4 is represented by the 3, because mean 2 and 4 belong to set 3.

4. $\mu_5 = \mu_3 > \mu_1 > \mu_2 = \mu_4$

   Likewise, this hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 -1 3

5. $\mu_5 = \mu_3 > \mu_1 > \mu_2,\ \mu_4$

   This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 -3 0

Note that, mean 4 is a free parameters, that is, means which are not restricted at all. The free parameters are denoted by a '0' and no group numbers are assigned (directly or indirectly). Here, as in the previous two examples, mean 2 belongs to group 3 and mean 4 belongs to another group. However, this is not group 4. Another example is given next.

6. $\mu_5 = \mu_3,\ \mu_4,\ \mu_1 > \mu_2$

This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 4 1 2
(Order) Restrictions
1 1 0 2 -3

The 1s indicate the equality constraint between mean 5 and 3. Note that, as mentioned in the previous example, no group number is assigned to free parameters. So, mean 4 belongs to another group than all the other means, but no group number is assigned. A 0 is filled in. Mean 1 belong to the next group, that is, group 2. The -3 represents the inequality constraint ">" between mean 1 and 2. Note that mean 2 indirectly belongs to group 3.

7. $\mu_5 = \mu_3 > \mu_1 < \mu_2 = \mu_4$

   This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 -1 3

8. $\mu_5 = \mu_3 > \mu_1 = \mu_2 > \mu_4$

   This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 2 -3

The 1s indicate the equality constraints between mean 5 and 3, the -3s represents the inequality constraint ">" between mean 3 and 1 and between 2 and 4. Note that mean 1 implicitly belongs to set 2. Therefore, the equality constraint between mean 1 and 2 is represented by the 2.

In the program *ConfirmatoryANOVA.exe* error messages are built in to detect wrongly stated hypotheses. But sometimes wrongly stated hypotheses are not detected, because the stated hypothesis represents another existing hypothesis. When accidently a 3 is given instead of a 2, another existing hypothesis is stated. Namely, the restriction "1 1 -3 3 -3" represents the hypothesis $\mu_5 = \mu_3 > \mu_1,\ \mu_2 > \mu_4$. So, care must be taken in writing down the hypothesis of interest.

9. $\mu_5 = \mu_3 > \mu_1,\ \mu_2 > \mu_4$

   This hypothesis can be represented by the following ordering of the group numbers and corresponding restriction:

Ordering of means in restriction
5 3 1 2 4
(Order) Restrictions
1 1 -3 3 -3

Equalities and About Equalities in BMS

In the $\bar{F}$-test, in ORIC and in BMS for strict equalities ($\delta = 0$), the two hypotheses $\mu_1 = \mu_2 = \mu_3$ (specified by 1 restriction, with ordering of means 1 2 3, and (order) restrictions 1 1 1) and $\mu_1 = \mu_2$, $\mu_2 = \mu_3$ (specified by 2 restrictions, both with ordering of means 1 2 3, and (order) restrictions 1 1 0 and 0 1 1) are equivalent.

However, the BMS approach in the program also provides the option to specify about equality constraints ($\delta > 0$). In that case, the second hypothesis is evaluated using $|\mu_1 - \mu_2| < \delta$ and $|\mu_2 - \mu_3| < \delta$, whereas the first hypothesis adds a third constraint: $|\mu_1 - \mu_3| < \delta$. The results of the first and second hypothesis may differ and therefore careful consideration of the formulation of hypotheses is important.

Set the Seed Value and Number of Iterations

The calculation of the p-value of the $\bar{F}$ and the penalty of the ORIC (i.e., PT) and BMS are sampling based approaches. For example, when generating data from a normal distribution (in order to determine the p-value of the $\bar{F}$

or the penalty of the ORIC), a seed value is needed. When using the same seed value, the same data will be "sampled". When looking at another seed value in a rerun of the same problem, one can also see how stable the results are. Thus, the p-value of the $\bar{F}$, the penalty of the ORIC (i.e., PT), and the results of BMS can differ for various seed values.

In case a result is not stable, the number of iterations needs be set higher. In the $\bar{F}$ test, the p-value depends on the number of iterations R_p. When using the ORIC, the penalty is dependent on the number of iterations R_{PT}. When doing BMS, the Gibbs sampler is used, which is based on a minimum of R_{BMS} iterations. These values can also be set in the input, namely in the second line of Input.txt (see Section **??**). The default values of the number of iterations in each method are: $R_p = 100,000$, $R_{PT} = 100,000$, and $R_{BMS} = 500,000$.

Note that the higher the number of iterations the higher the computing time. If one lowers the number of iterations (in order to lower the computing time), one must be aware that this probably affects the stability of the results. Furthermore, when the initial number of iterations for BMS (i.e., R_{BMS}) is lowered, the computing time is not necessarily decreased, because of the requirement of a minimum of 100 prior hits.

Error Messages

In the program *ConfirmatoryANOVA.exe*, error messages are built in to detect wrongly stated hypotheses. However, it does not detect all wrongly stated hypotheses, since the stated hypothesis can represent another existing hypothesis, as is made clear in the previous section.

It is also possible to state other input wrongly. For example, the wrong number of restrictions is given. When making a mistake, an informative warning will be given.

It is, however, possible to make a mistake that we have not foreseen. In that case, check the input in *Input.txt* and compare it to the data. If you cannot solve the problem, send the input and data file to r.m.kuiper@uu.nl.

Modification of the Second Half of ***Input.txt***

For all three methods (i.e., the $\bar{F}$ test, ORIC, and BMS), all hypotheses of interest must be given explicitly. In case BMS is performed, two additional specifications need to be made: The desired δ (for exact equalities specify $\delta = 0$, for an about equality any positive number can be specified) and the prior vagueness pv (default recommendation $pv = 2$; any positive number may be specified). This must be done in the second half of *Input.txt*.

If the hypotheses of interest are the set of hypotheses specified in Lucas (2003), that is,

$$
\begin{aligned}
H_0: &\quad \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5, \\
H_1: &\quad \mu_5 \ = \mu_3 > \mu_1 > \mu_2,\ \mu_3 > \mu_4 > \mu_2, \\
H_2: &\quad \mu_3 > \mu_1 > \mu_4 \ = \mu_5 > \mu_2, \\
H_A: &\quad \mu_1,\ \mu_2,\ \mu_3,\ \mu_4,\ \mu_5,
\end{aligned} \tag{2}
$$

Input.txt has the following format (where “$< \ldots >$” is not part of the format, but is used to give remarks):

```
Number of models to be compared
<Fill in the number of models / hypotheses you want to compare; e.g.,>
4
Number of restrictions per model
<Fill in, for every model / hypothesis, the number of restrictions that
represent that model / hypothesis; e.g.,>
1
2
1
1
Ordering of means in restriction
<Fill in the ordering of the means / group numbers for each restriction
for every model / hypothesis. The orderings per restrictions are separated
by an "enter". The ordering consists of the numbers 1 to "the total
number of groups". For more details see "Basic Elements of Writing
Constraints" and "Combinations of Basic Elements"; e.g.,>
1 2 3 4 5
5 3 1 2 4
3 4 2 1 5
3 1 4 5 2
1 2 3 4 5
(Order) Restrictions
<Fill in the restrictions. This must be done in a certain manner, which
is explained above in "Basic Elements of Writing Constraints" and
"Combinations of Basic Elements"; e.g.,>
1  1  1  1  1
1  1 -3 -3  0
1 -3 -3  0  0
1 -3 -3  3 -3
0  0  0  0  0
When BMS is performed, an interval for equality relations (delta) is
needed and a parameter for prior vagueness (pv)
<Fill in δ ≥ 0 and pv > 0; e.g.,>
0.0 2.0
```

Note that the numbering of the hypotheses is 1 to “the number of models”, which is not the same as the numbering in (2). Thus, the following models/hypotheses will be analyzed in *ConfirmatoryANOVA.exe*:

$$\begin{aligned}
H_1: &\quad \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5, \\
H_2: &\quad \mu_5 = \mu_3 > \mu_1 > \mu_2,\ \mu_4, \\
&\quad \mu_3 > \mu_4 > \mu_2,\ \mu_1,\ \mu_5, \\
H_3: &\quad \mu_3 > \mu_1 > \mu_4 = \mu_5 > \mu_2, \\
H_4: &\quad \mu_1,\ \mu_2,\ \mu_3,\ \mu_4,\ \mu_5.
\end{aligned}$$

Save and close

When you have modified *Input.txt* (such that it applies to your data), you should save and close it.

Run *ConfirmatoryANOVA.exe*

When *ConfirmatoryANOVA.exe* is done, the output file *Output.txt* will be created in the folder you are working in.

Output.txt

Output.txt gives the results of the requested analyses. In case of the example of *Input.txt*, *Output.txt* will look as follows (provided that all the analyses, that is, SWFq and TK, PCIC, exploratory BMS, $\bar{F}$, ORIC, and confirmatory BMS, are performed):

This program is free of use. However, when results obtained with this program are published, please refer to:

Rebecca M. Kuiper, Irene Klugkist, and Herbert Hoijtink (2010).
A Fortran 90 Program for Confirmatory Analysis of Variance.
Journal of Statistical Software, 34(8), 1-31.
URL http://www.jstatsoft.org/v34/i08/.
N.B. This paper is available upon request (R.M.Kuiper@uu.nl).

Summary of observed data

Group number, means, standard deviations, and sample sizes per group

1	**2.33**	**1.86**	**30**
2	**1.33**	**1.15**	**30**
3	**3.20**	**1.79**	**30**
4	**2.23**	**1.45**	**30**
5	**3.23**	**1.50**	**30**

Restricted means

Group number:	1	2	3	4	5
Sample means:	**2.33**	**1.33**	**3.20**	**2.23**	**3.23**
Hypothesis 1	**2.46**	**2.46**	**2.46**	**2.46**	**2.46**
Hypothesis 2	**2.33**	**1.33**	**3.20**	**2.23**	**3.23**
Hypothesis 3	**2.33**	**1.33**	**3.21**	**2.23**	**3.21**
Hypothesis 4	**2.60**	**1.33**	**3.20**	**2.60**	**2.60**

– Fbar test –

Results of the Fbar test for the null hypothesis **1** and the unconstrained hypothesis **4**

Hypotheses numbers	Fbar value	p-value
1 versus **4**	**30.27**	**0.00**

Results of the "ordered alternative" Fbar test

Ordered-hypothesis number	Fbar value	p-value
H0 versus **2**	**30.26**	**0.00**
H0 versus **3**	**22.91**	**0.00**

Results of the "ordered null" Fbar test

Ordered-hypothesis number	Fbar value	p-value
2 versus Ha	**0.01**	**1.00**
3 versus Ha	**7.36**	**0.07**

Resisual sum of squares

Hypothesis 1 **432.53**
Hypothesis 2 **357.85**
Hypothesis 3 **375.99**
Hypothesis 4 **0.00**

– ORIC –

The value of the Order-Restricted Information Criterion (ORIC) =
-2 * log likelihood + 2 * penalty:

for Hypothesis 1, ORIC = -2 * **-292.27** + 2 * **2.00** = **588.54**
for Hypothesis 2, ORIC = -2 * **-278.05** + 2 * **3.20** = **562.49**
for Hypothesis 3, ORIC = -2 * **-281.76** + 2 * **3.14** = **569.80**
for Hypothesis 4, ORIC = -2 * **-278.05** + 2 * **6.00** = **568.10**

The preferred hypothesis, according to the Order-Restricted Information Criterion, of the hypotheses to be compared is hypothesis number **2**,
with the following ordering(s) of means:
5 3 1 2 4
3 4 2 1 5

and corresponding restriction(s):
1 1 -3 -3 0
1 -3 -3 0 0

– BMS –

The resulting Bayes factor (BF) values (of the order-restricted hypotheses versus the unconstrained hypothesis) and the posterior model probabilities (PMP) of the order-restricted hypotheses with respect to the whole set of hypotheses:

	BF	PMP
Hypothesis 1	**0.00**	**0.00**
Hypothesis 2	**66.21**	**0.96**
Hypothesis 3	**1.62**	**0.02**
Hypothesis 4	**1.00**	**0.01**

The preferred hypothesis, according to confirmatory Bayesian model selection, of the hypotheses to be compared is hypothesis number **2**, with the following ordering(s) of means:
5 3 1 2 4
3 4 2 1 5
and corresponding restriction(s):
1 1 -3 -3 0
1 -3 -3 0 0

Specification of the encompassing prior:

For all means, the same normal prior with mean
2.28
and variance
2.32
is used.

For the residual variance, a scaled inverse chi-square with degrees of freedom
1.00
and scale parameter
2.50
is used.

The parts which are in bold are the parts, which will change per data set and/or set of hypotheses. More information about the output of the methods can be found in Kuiper and Hoijtink (2010) and Kuiper et al. (2010).

When using the ORIC the preferred model/hypothesis is given, in this example it is "hypothesis 2". Besides that, the corresponding ordering(s) of means and restriction(s) are given, which are given in the *Input.txt*. Thus, according to the ORIC, $H_2 : \mu_5 \ = \mu_3 > \mu_1 > \mu_2,\ \mu_3 > \mu_4 > \mu_2$ (i.e., H_1 from (2)) is the preferred hypothesis. The same holds for BMS.

References

Klugkist, I., Laudy, O. and Hoijtink, H. (2005). Inequality Constrained Analysis of Variance: A Bayesian Approach. *Psychological methods, 10*, 477–493.

Kuiper, R. M. and Hoijtink, H. (2010). Comparisons of Means Using Exploratory and Confirmatory Approaches. *Psychological Methods, 15(1)*, 69-86.

Kuiper, R. M., Klugkist, I., and Hoijtink, H. (2010). A Fortran 90 Program for Confirmatory Analysis of Variance. *Journal of Statistical Software, 34(8)*, 1-31.

Lucas, J.W. (2003). Status processes and the institutionalization of women as leaders. *American Sociological Review, 68*, 464-480.