Pop-Up Menus in FAiR

Ben Goodrich

February 1, 2008

This vignette is intended to be a quick reference that defines terms and illustrates how to use the pop-up menus. Large parts will not make a great deal of sense outside the introduction to SEFA and lexical optimization in Goodrich (2008a) and its (unwritten) extension to EFA in Goodrich (2008b).

Contents

1 Notation

Some notation is provided so that the reader can refer back to it. For simplicity, I do not distinguish the notation for a population parameter from the notation for an estimate of that population parameter here.

- n is the number of outcome variables
- $\mathbf{S}$ is the $n \times n$ correlation matrix among outcomes in the sample
- $r_1 \geq 1$ is the number of factors (at the first level)
- $\boldsymbol{\beta}$ is the $n \times r_1$ primary pattern matrix (at the first level)
- $\boldsymbol{\Phi}$ is the $r_1 \times r_1$ correlation matrix among primary factors (at the first level)
- $\boldsymbol{\Theta}^2$ is the $n \times n$ diagonal covariance matrix among unique factors (at the first level)
- $\mathbb{E}[\mathbf{S}] = \mathbf{C} = \boldsymbol{\beta}\boldsymbol{\Phi}\boldsymbol{\beta}' + \boldsymbol{\Theta}^2$ is the $n \times n$ expectation of $\mathbf{S}$ as a function of parameters (at the first level)
- $\boldsymbol{\Pi} = \boldsymbol{\beta}\boldsymbol{\Phi}$ is the $n \times r_1$ primary structure matrix (at the first level)
- $\boldsymbol{\Gamma} = \boldsymbol{\beta} \times \boldsymbol{\Pi}$ is the $n \times r_1$ factor contribution matrix (at the first level) where $\times$ indicates element-by-element multiplication rather than matrix multiplication
- $\mathbf{D}^{-2} = \text{Diag}\left(\boldsymbol{\Phi}^{-1}\right)$ is the $r_1 \times r_1$ diagonal of the inverse of the correlation matrix among primary factors (at the first level)
- $\mathbf{D} = \left[\mathbf{D}^{-2}\right]^{-\frac{1}{2}}$ is the $r_1 \times r_1$ correlation matrix between primary and reference factors (at the first level)
- $\boldsymbol{\Psi} = \mathbf{D}\boldsymbol{\Phi}^{-1}\mathbf{D}$ is the $r_1 \times r_1$ correlation matrix among reference factors (at the first level)
- $\boldsymbol{\Upsilon} = \boldsymbol{\beta}\mathbf{D}$ is the $n \times r_1$ reference structure matrix (at the first level)
- $r_2 \geq 0$ is the number of second-order factors
- $\boldsymbol{\Delta}$ is the $r_1 \times r_2$ second-order primary pattern matrix
- $\boldsymbol{\Xi}$ is the $r_2 \times r_2$ correlation matrix among second-order primary factors
- $\boldsymbol{\Omega}^2$ is the $r_1 \times r_1$ diagonal covariance matrix among second-order unique factors
- $\boldsymbol{\Phi} = \boldsymbol{\Delta}\boldsymbol{\Xi}\boldsymbol{\Delta}' + \boldsymbol{\Omega}^2$ is the $r_1 \times r_1$ second-order equation for the primary factors as a function of second-order parameters
- $\digamma = \boldsymbol{\Delta}\boldsymbol{\Xi}$ is the $r_2 \times r_1$ primary structure matrix at the second level
- $\eth = \boldsymbol{\Delta} \times \digamma$ is the $r_2 \times r_1$ factor contribution matrix at the second level where $\times$ indicates element-by-element multiplication rather than matrix multiplication
- $\mathbf{Z}^{-2} = \text{Diag}\left(\boldsymbol{\Xi}^{-1}\right)$ is the $r_2 \times r_2$ diagonal of the inverse of the correlation matrix among primary factors at the second level
- $\mathbf{Z} = \left[\mathbf{Z}^{-2}\right]^{-\frac{1}{2}}$ is the $r_2 \times r_2$ correlation matrix between primary and reference factors at the second level
- $\mho = \boldsymbol{\Delta}\mathbf{Z}$ is the $r_1 \times r_2$ reference structure matrix at the second level

2 SEFA or CFA via `Factanal()`

To start this sequence, I typed the command

```
mental.tests <- Factanal(covmat = Harman74.cor, factors = 5, model = "SEFA")
```

2.1 Simultaneous Second-Order Model

If $r_1 \geq 3$, the user will be asked whether to estimate a simultaneous second-order model, which decomposes the correlation matrix among first-order primary factors as a function of fewer second-order factors. For example, some people define "general intelligence" to be the second-order factor that drives various first-order (primary) mental abilities. If a second-order model is estimated, and if $r_1 \geq 5$, the user will be asked how many second-order factors (r_2) to extract.

If a second-order model is specified, the tasks are to set bounds on the correlations among second-order factors (if $r_2 \geq 2$), to fix values of the second-order coefficients (for CFA and mixed SEFA models), to set bounds on the second-order coefficients, and to choose a mapping rule for the second-order coefficients (if $r_2 \geq 2$ in a SEFA model).

2.1.1 Deciding on a Simultaneous Second-Order Model

Figure 1 shows two dialog boxes. First, the user is asked whether to estimate a simultaneous second-order model. If the user were to choose "No" at this point, there would be no second-order model and the user would then be asked about bounds on the off-diagonals of $\mathbf{\Phi}$. Such a dialog would be similar to that in section 2.1.2, except that "level 2" would be replaced by "level 1". Then the dialog would cut directly to that described in section 2.2.

For the sake of completeness, I will assume that the user answered "Yes" to the question of whether a simultaneous second-order model should be estimated. At this point, the dialog asks the user to specify the value of r_2. If $r_1 = 3$ or $r_1 = 4$, then r_2 can only be one. If $r_1 \geq 5$, then r_2 can be two. If $r_1 \geq 6$, then r_2 can be three. Rarely would one need to estimate a model with larger numbers for r_1 and r_2. I will assume that the user has chosen $r_2 = 2$.

2.1.2 Bounds on the Correlations Among the Primary Factors $(\mathbf{\Xi})$ at Level 2

The next menu asks about bounds on the off-diagonals of $\mathbf{\Xi}$, as shown in figure 2. If $r_2 \leq 1$, this question would not be asked.

Since $\mathbf{\Xi}$ is constrained to be a positive definite correlation matrix, the off-diagonals cannot exceed ± 1. It is possible, but often not necessary, to impose stricter bounds. Occasionally, there can be optimization difficulties if any correlation gets too close to ± 1, and in that case the user may want to narrow the acceptable interval to $(-0.9, 0.9)$ or so.

If "No" is chosen, the dialog will skip to section 2.1.3 using the $(-1, 1)$ interval for all correlations among second-order factors. For the sake of completeness, I will assume the user answers "Yes" to this question.

Since $\mathbf{\Xi}$ is a correlation matrix, the diagonal elements are constrained to be 1.0 and the diagonal elements will be ignored if they are changed. Since a correlation matrix is symmetric, it is not necessary to change the values below the diagonal and they will be ignored regardless. It is only necessary to specify the values above the diagonal. In this case, I have specified a lower bound of -0.9 on the only correlation among the two primary factors at level 2, which is a very conservative choice.

When satisfied with the lower-bounds on the off-diagonals of $\mathbf{\Xi}$, click the `x` at the top right (or "Quit", but "Quit" does not appear in the Windows version). Once `x` is pressed, the process is repeated for the upper bounds on the off-diagonals of $\mathbf{\Xi}$.

Figure 1: Simultaneous Second-Order Model with $r_2 = 2$

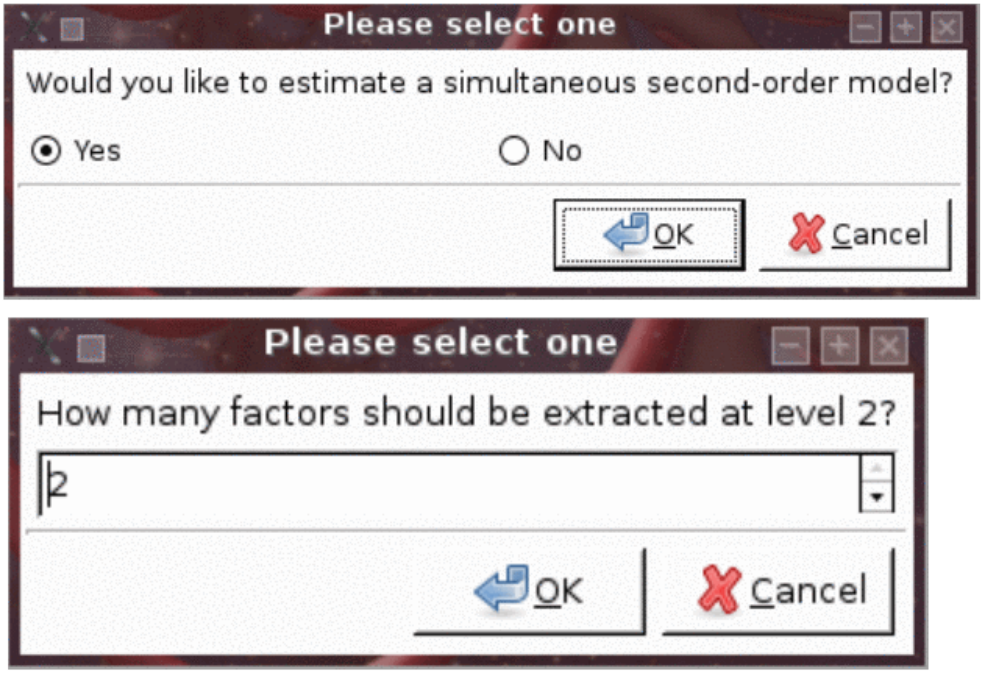


Figure 2: Slightly Limiting the Bounds on the Correlation Between Second-Order Primary Factors

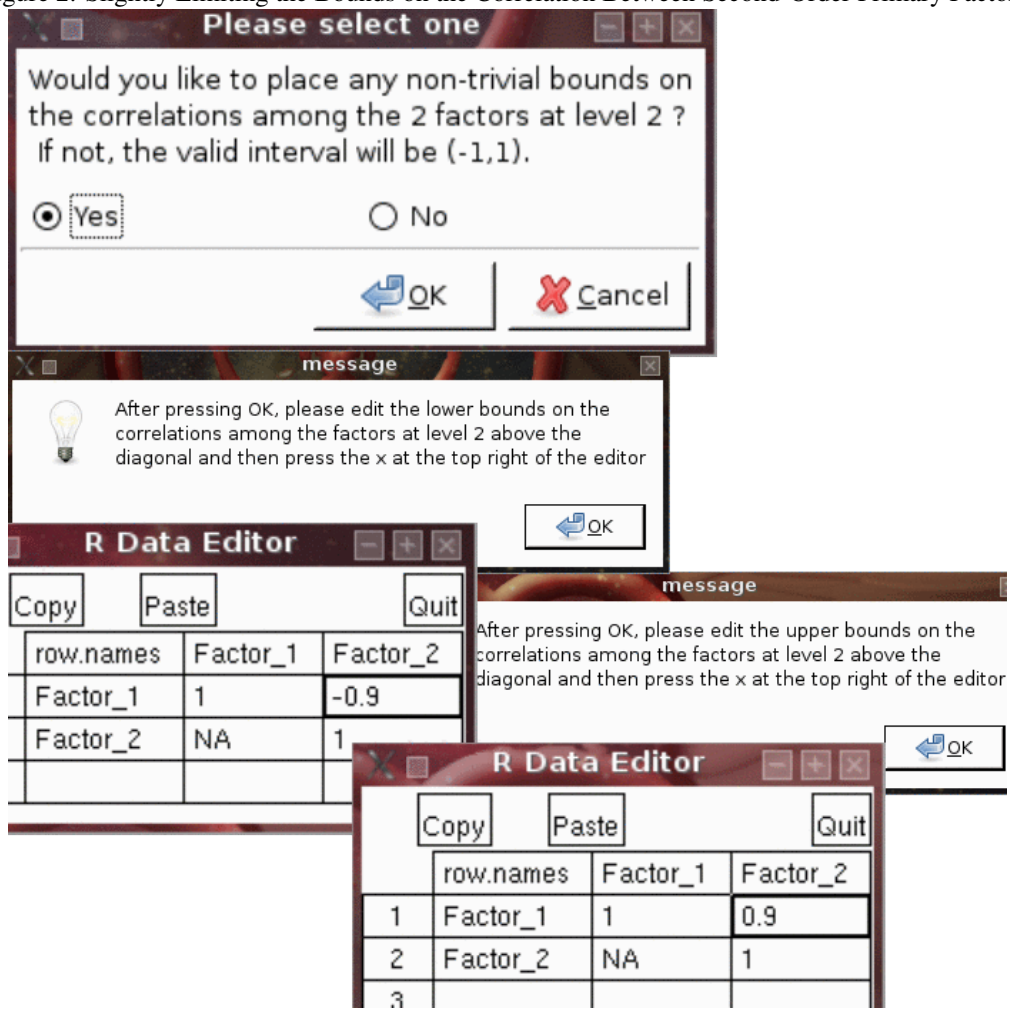


Figure 3: (Not) Pegging Coefficients to Specified Values

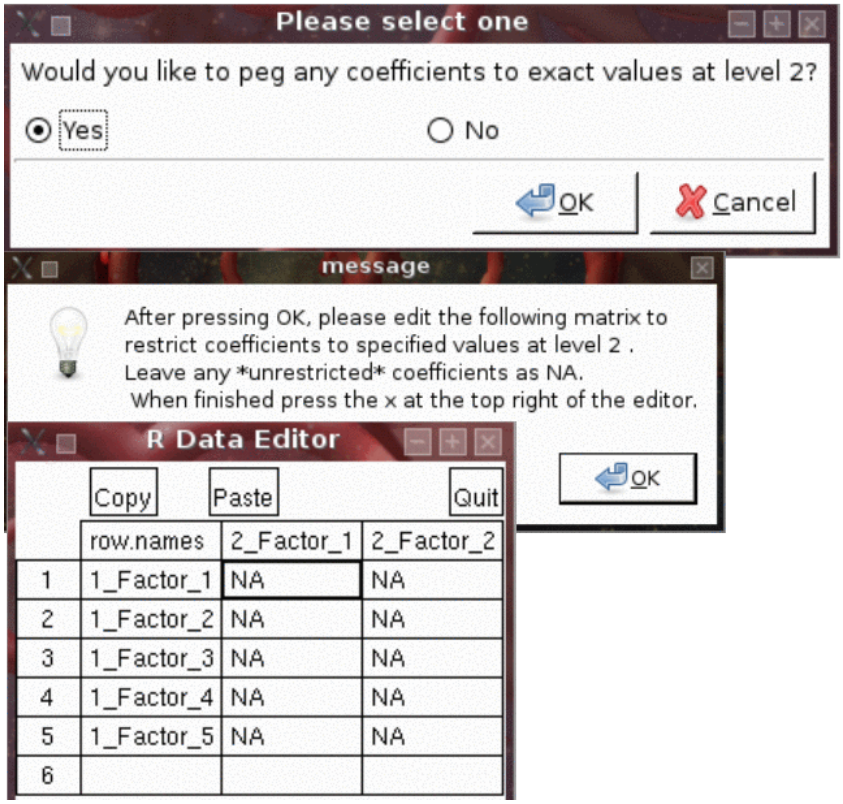
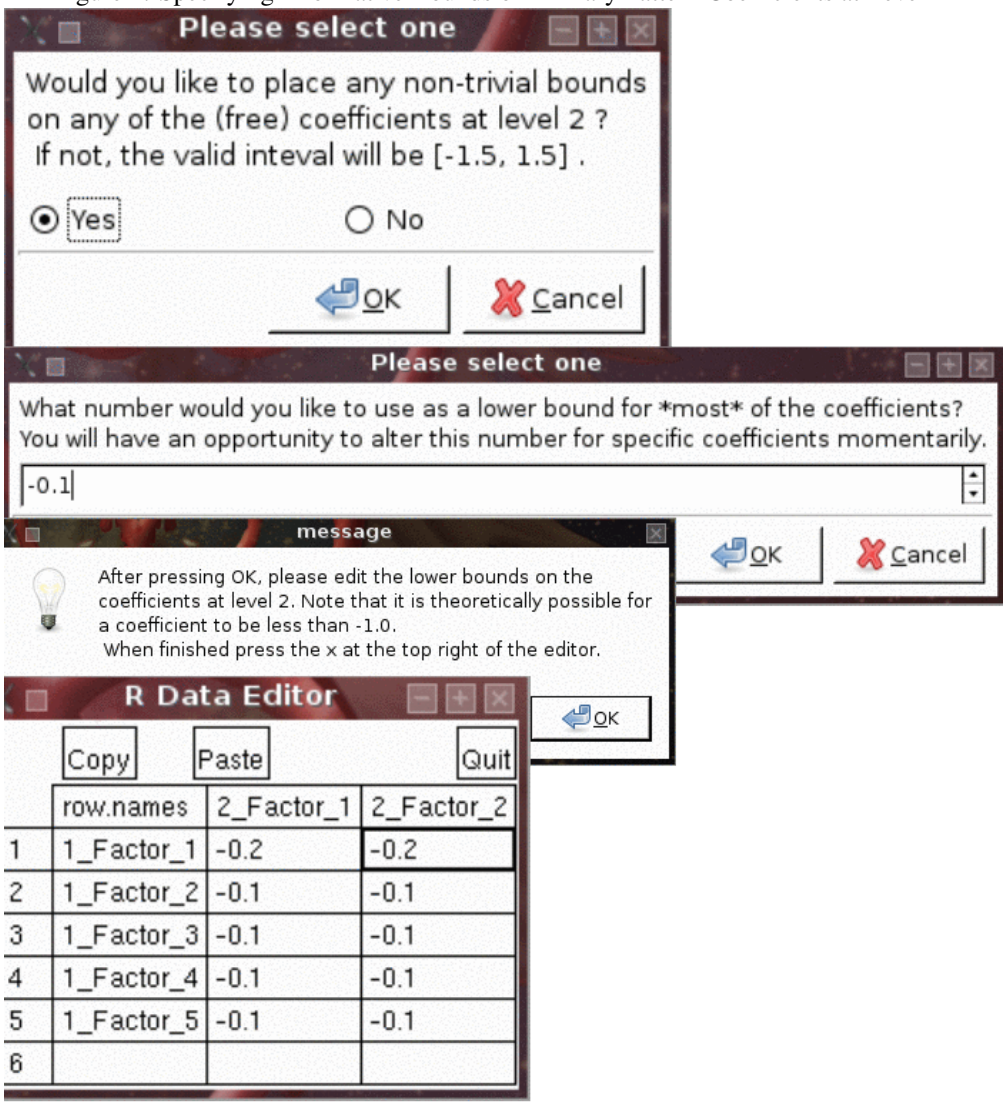


Figure 4: Specifying Informative Bounds on Primary Pattern Coefficients at Level 2



2.1.3 Specifying Values of the Primary Pattern Matrix ($\boldsymbol{\Delta}$) at Level 2

Since this example is a SEFA, the next dialog (shown in figure 3) asks the user whether any elements of $\boldsymbol{\Delta}$ should be pegged to specific values. If a CFA model were estimated, this question would not appear because it is obligatory to peg some elements of $\boldsymbol{\Delta}$ in a CFA model. In a SEFA model, it is possible to answer "No" to this question in order to estimate a "pure" SEFA model. However, for illustrative purposes, I will assume that the user answers "Yes" to estimate a "mixed SEFA" model.

At this point, the user should change the appropriate cells of $\boldsymbol{\Delta}$ from NA to numbers. For example, the user might change the first row to $\begin{bmatrix}1.0 & 0.0\end{bmatrix}$ to make the first factor at the second level collinear with the first factor at the first level. Any elements of $\boldsymbol{\Delta}$ that are to be *unrestricted* should be left as NA.

If a CFA model were being estimated, it would be necessary to specify restrictions on $\boldsymbol{\Delta}$ to satisfy the theorem on rotational indeterminancy in Howe (1955). Namely there should be at least $r_2 - 1$ zeros in each column of $\boldsymbol{\Delta}$ such that all r_2 submatrices of $\boldsymbol{\Delta}$ with zeros in the pth column of $\boldsymbol{\Delta}$ are of rank $r_2 - 1$.

In this case, a mixed SEFA model is being estimated, so it is possible to specify fewer cells of $\boldsymbol{\Delta}$ including no restricted cells at all. Here I assume the user has changed his or her mind about restrictions on $\boldsymbol{\Delta}$ and now wants all cells to be unrestricted. Thus, all cells remain NA.

2.1.4 Bounds on Cells of the Primary Pattern Matrix ($\boldsymbol{\Delta}$) at Level 2

The next question is what bounds should be placed on the cells of $\boldsymbol{\Delta}$. The dialog shown in figure 4 is similar to that in section 2.1.2, except that the bounds are being placed on the cells of $\boldsymbol{\Delta}$ rather than the off-diagonals of $\boldsymbol{\Xi}$. It is always possible to select "No", but doing so does not make the cells of $\boldsymbol{\Delta}$ unbounded. Rather, fairly wide bounds are used, in particular ± 1.5. In some cases, it may be necessary to specify wider bounds, but in most cases if the user answers "Yes", it would be to narrow the bounds.

If the user specifies bounds, the number chosen in the next dialog fills every free cell of the subsequent editable menu. The fixed coefficients, if any, would be filled with their values. The user can then edit particular cells as necessary. Here I have changed both cells in the first row to -0.2 to permit a more diffuse positive manifold for the first second-order factor.

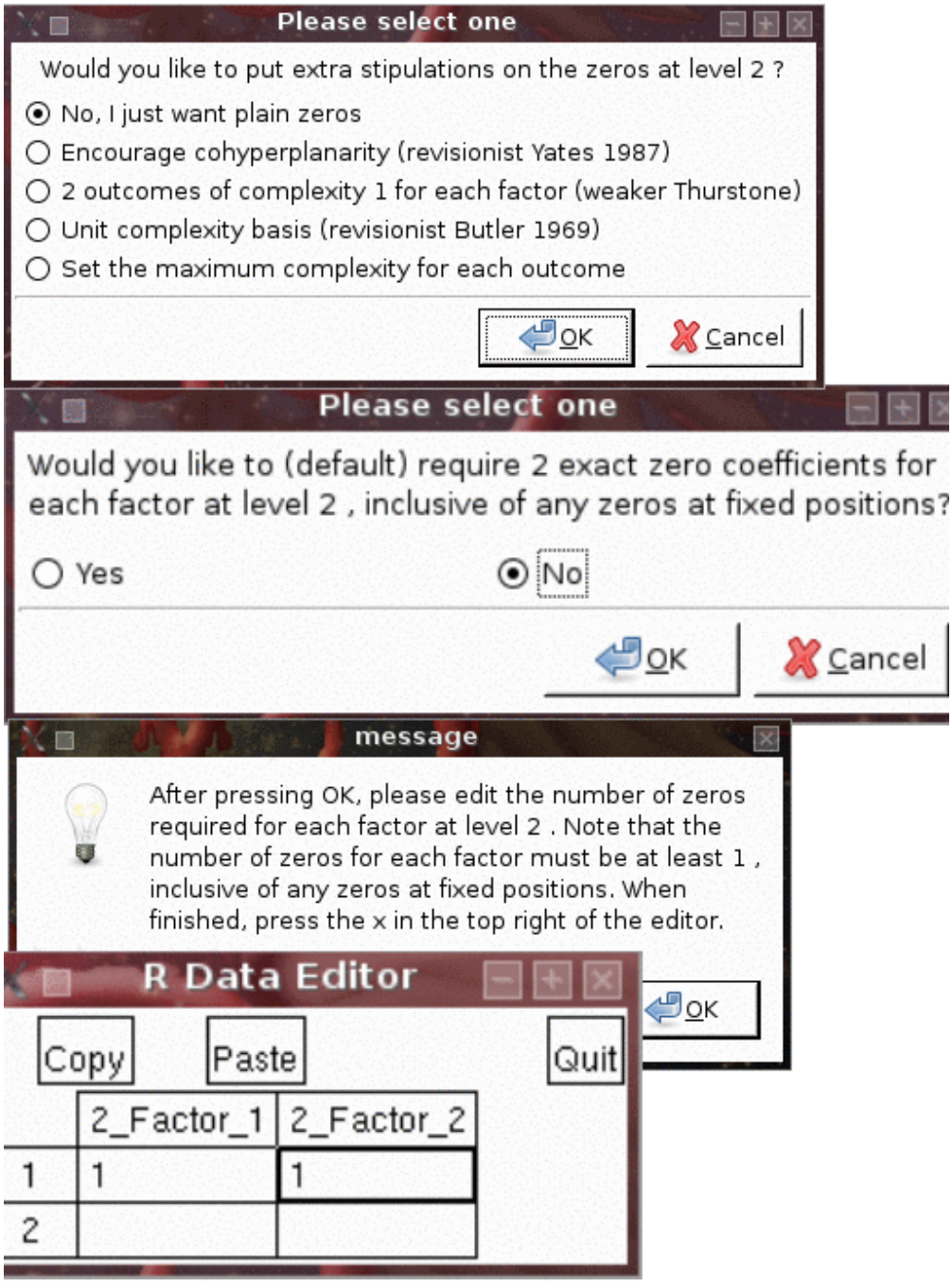
It should be emphasized that it is possible for coefficients to be less than -1.0 when $r_2 > 1$. The decision here to limit the cells of $\boldsymbol{\Delta}$ to be greater than -0.2 or -0.1 is a *substantive* constraint that can be wrong and can be tested. If strong bounds are used, it may be best to specify `boundary.enforcement = 2` in the call to `Factanal`.

When finished, press `x` (or "Quit" but "Quit" does not appear on the Windows version), to repeat this procedure for the upper bounds on the cells of $\boldsymbol{\Delta}$. It is necessary that some of the valid intervals for the cells of $\boldsymbol{\Delta}$ include zero so that the Howe (1955) theorem can be satisfied in a SEFA model. It is difficult to generally prove when inequality restrictions on coefficients are sufficient to eliminate rotational indeterminacy, so the Howe (1955) theorem is recommended.

2.1.5 Mapping Rules for the Primary Pattern Matrix ($\boldsymbol{\Delta}$) at Level 2

In the SEFA model presumed here, the questions shown in figure 5 are very important and pertain to the mapping rule used to squash certain cells of $\boldsymbol{\Delta}$ to zero. If a CFA model were being estimated, this question would not appear because the Howe (1955) theorem must be satisfied by *a priori* restrictions. Let b_p be the number of zeros required for the pth column of $\boldsymbol{\Delta}$. Most of the mapping rules at the second level are defined on the reference structure matrix $\left(\check{\boldsymbol{\Upsilon}}\right)$ and the factor contribution matrix $\left(\check{\boldsymbol{\Im}}\right)$, rather than the primary pattern matrix ($\check{\boldsymbol{\Delta}}$), where the supra-check indicates a prelimary matrix that is constructed by filling the free elments with $\boldsymbol{\theta}$ (in symbols) or `par` (in R syntax). By "smallest", I always mean in absolute value. The distinctions between mapping rules are fairly minor in this case when $r_2 = 2$, but are defined in general as follows:

Figure 5: Mapping Rules at Level 2 in a SEFA



- "No, I just want plain zeros". This mapping rule simply squashes the b_p smallest elements in the pth column of $\check{\mathcal{O}}$ to zero, implying that $\boldsymbol{\Delta}_p$ has b_p zeros.
- "Encourage cohyperplanrity (revisionist Yates 1987)". This mapping rule squashes $r_2 - 1$ elements in the pth column of $\boldsymbol{\Delta}$ using the following piecewise function $\forall q \neq p$:

$$\Delta_{jp} \quad = \quad \begin{cases} 0 & \text{if } \check{\mathcal{O}}_{jq} - \check{\mathcal{O}}_{jq} = \max\left\{\check{\mathcal{O}}_q - \check{\mathcal{O}}_p\right\} \\ \check{\Delta}_{jp} & \text{otherwise,} \end{cases}$$

  which tends to place a zero in the pth column of $\boldsymbol{\Delta}$ at the row where the factor contribution of some other column is large for each of the remaining columns.
- "r_2 'outcomes' (really primary factors) of complexity $r_2 - 1$ for each factor (weaker Thurstone)". This mapping rule is similar to the "plain zeros" mapping rule in that r_2 small elements in the pth column of $\check{\mathcal{O}}$ are squashed to zero, except that there is an additional stipulation that no "outcome" (really primary factor) is of complexity less than $r_2 - 1$ where the complexity of the jth "outcome" (primary factor) is the number of non-zeros in $\mathcal{O}_j$. However, any "non-zero" coefficient can be arbitrarily close to zero. Thus, each column of $\boldsymbol{\Delta}$ will have r_2 *exact* zeros.
- "Unit complexity basis (revisionist Butler 1969)". This mapping rule uses the dishinguishability weights defined in Yates (1987) to select r_2 primary factors that are of unit complexity. In particular, let

$$\check{w}_p \quad = \quad \frac{1}{\sum \check{\mathcal{O}}_{\bullet\bullet}} \sum_{j=1}^{r_1} \check{\mathcal{O}}_{jp}$$

  where $\sum \check{\mathcal{O}}_{\bullet\bullet} = \sum_{j=1}^{r_1} \sum_{p=1}^{r_2} \check{\mathcal{O}}_{jp}$ is the weight applied the the pth second-order factor such that $\sum_{p=1}^{r_2} w_p = 1$. This equation is similar to equation 118 in Yates (1987, p. 145). Then let

$$\check{c}_j \quad = \quad \frac{\sum_{p=1}^{r_2} \check{w}_p \check{\Delta}_{jp}}{\sum_{p=1}^{r_2} \check{\mathcal{O}}_{jp}}$$

  be the directional cosine between $\check{\boldsymbol{\Delta}}_j$ and the (weighted) centroid of $\check{\boldsymbol{\Delta}}$, which is the same as equation 119 in Yates (1987, p. 145). Finally, let

$$d_j^w \quad = \quad \min\left\{1, \frac{(r_2)^2}{r_2 - 1} \check{c}_j \left(\check{c}_j - 1\right)\right\}$$

  be the distinguishability weight for $\check{\boldsymbol{\Delta}}_j$, which is the same as equation $78a$ in Yates (1987, p. 112). Then, $\forall p$ choose a first-order factor to be of unit complexity among first-order factors where $\check{\mathcal{O}}_{jp} = \max\left\{\check{\mathcal{O}}_j\right\}$ such that d_j^w is maximal. For such first-order factors, squash $\boldsymbol{\Delta}_j$ to zero $\forall q \neq p$. At the end of these steps, there will be r_2 first-order factors that are collinear with the second-order primary axes and there will be $r_2 - 1$ zeros in each column of $\boldsymbol{\Delta}$.
- "Set the maximum complexity for each outcome." This mapping rule is similar to the "plain zeros" mapping rule, except that it is applied to each *row* of $\boldsymbol{\Delta}$ rather than the columns. In general, $\forall j$, the user can choose a number between 1 and r_2 for the maximum number of non-zero coefficients in $\boldsymbol{\Delta}_j$. More details on this mapping rule are given in section X.

In some cases, b_p may be greater than the number of zeros created by whichever mapping rule is chosen. In that case, the remaining zeros are filled in by squashing the b_p smallest elements in the pth column of the *recalculated* reference structure matrix after the mapping rule has been applied and some reference structure correlations are zero. Thus, it is necessary to specify b_p, and the dialog box for doing so is also shown in figure 5. By default, the user is asked whether to set $b_p = r_2 \, \forall p$, which can be accepted by answering "Yes" to the following question but here I answer "No" to specify $b_p = r_2 - 1 \, \forall p$.

This concludes the dialog boxes for the second-order model.

2.2 First-Order Model

The tasks are to set bounds on the correlations among first-order factors (if a second-order model is not estimated), to fix values of the first-order coefficients (for CFA and mixed SEFA models), to set bounds on the first-order coefficients, and to choose a mapping rule for the first-order coefficients (if $r_2 \geq 2$ in a SEFA model).

Note that if there is no second-order model, the user will be asked about bounds on the off-diagonals of $\mathbf{\Phi}$. This dialog is similar to that in section 2.1.2, except that "level 2" is replaced by "level 1" in the prompts.

2.2.1 Specifying Values of the Primary Pattern Matrix $(\boldsymbol{\beta})$ at Level 1

Since this example is a SEFA, the dialog shown in figure 6 asks the user whether any elements of $\boldsymbol{\beta}$ should be pegged to specific values. If a CFA model were estimated, this question would not appear because it is obligatory to peg some elements of $\boldsymbol{\beta}$ in a CFA model. In a SEFA model, it is possible to answer "No" to this question to estimate a "pure" SEFA model. However, for illustrative purposes, I will assume that the user answers "Yes" to estimate a "mixed" SEFA model.

At this point, the user should change the appropriate cells of $\boldsymbol{\beta}$ from NA to numbers. For example, the user might change the first row to $\begin{bmatrix}\text{NA} & 0.0 & 0.0 & 0.0 & 0.0\end{bmatrix}$ to make the first factor collinear with visual perceptiveness. Any elements of $\boldsymbol{\beta}$ that are *unrestricted* should be left as NA. Continuing this example, by leaving the first factor unrestricted for the visual perception test, the user allows the test to have measurement error.

If a CFA model were being estimated, it would be necessary to specify restrictions on $\boldsymbol{\beta}$ to satisfy the theorem on rotational indeterminancy in Howe (1955). Namely there should be at least $r_1 - 1$ zeros in each column of $\boldsymbol{\beta}$ such that all r_1 submatrices of $\boldsymbol{\beta}$ with zeros in the pth column of $\boldsymbol{\beta}$ are of rank $r_1 - 1$.

In this case, a SEFA model is being estimated, so it is possible to specify fewer cells of $\boldsymbol{\beta}$ including no restricted cells at all. Here I assume the user has changed his or her mind about restrictions on $\boldsymbol{\beta}$ and now wants all cells to be unrestricted. Thus, all cells are left as NA.

2.2.2 Bounds on Cells of the Primary Pattern Matrix $(\boldsymbol{\beta})$ at Level 1

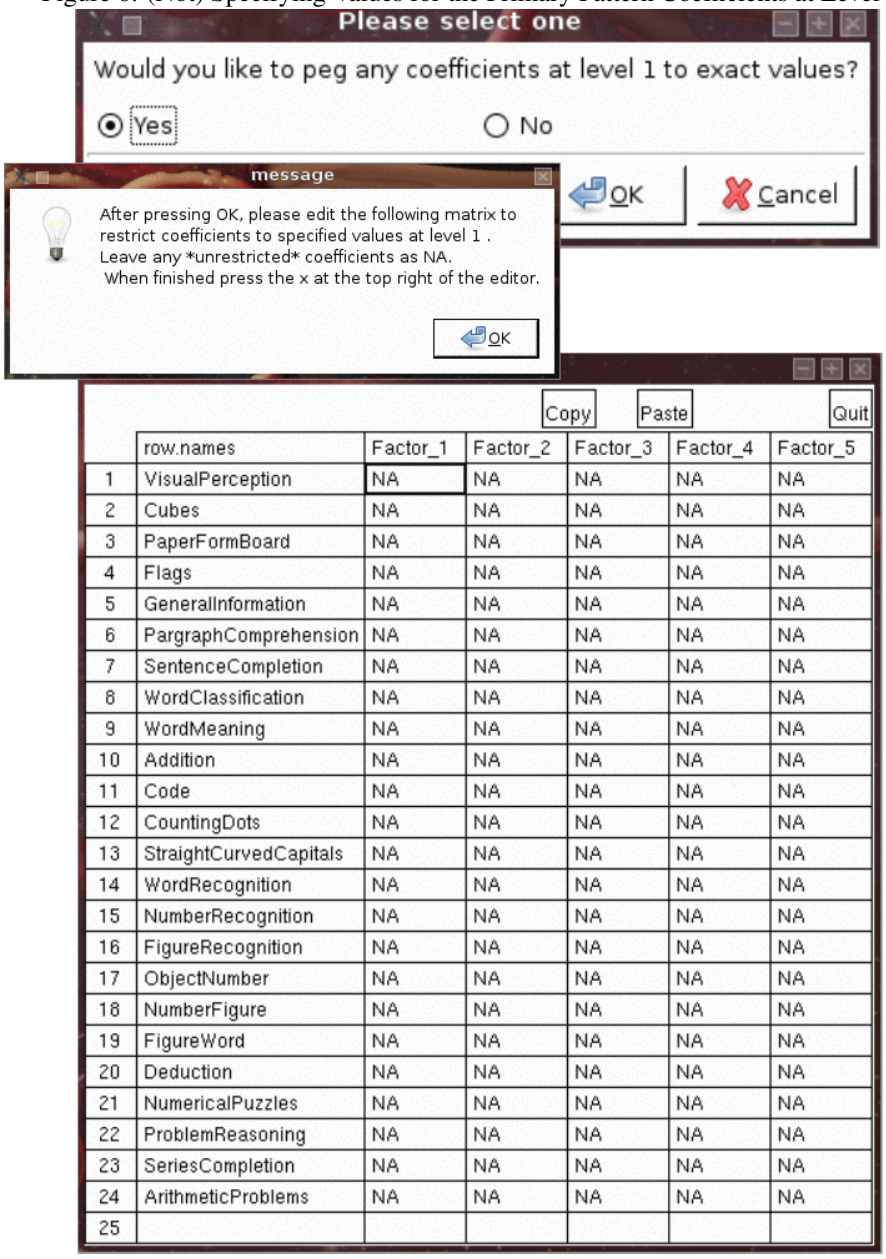
The next question is what bounds should be placed on the cells of $\boldsymbol{\beta}$. The dialog in figure is similar to that in section 2.1.4, except that the bounds are being placed on the cells of the primary pattern matrix at level 1 $(\boldsymbol{\beta})$ rather than the primary pattern matrix at level 2 $(\mathbf{\Delta})$. It should be kept in mind that the model is estimated on a *correlation* matrix of outcomes $(\mathbf{S})$ when choosing bounds. Thus, it is somewhat uncommon — but nevertheless possible — for a coefficient to be greater than ± 1. It is always possible to select "No", but doing so does not make the cells of $\boldsymbol{\beta}$ unbounded. Rather, fairly wide bounds are used, in particular ± 1.5. In some cases, it may be necessary to specify wider bounds, but in most cases if the user answers "Yes", it would be to narrow the bounds.

After an initial number is chosen to fill all the cells in the editor, the user can change some or all of the cells as necessary. For example, to estimate the model under a positive manifold restriction, the user could set all lower bounds to -0.2 or something. When finished with the lower bounds, click the $\mathbf{x}$, at which point the process is repeated for the upper bounds. It is important for some, if not all, of the valid intervals to include zero in a SEFA model so that the rotational indeterminancy can be eliminated by mapping some coefficients to zero.

2.2.3 Mapping Rules for Coefficients at Level 1

In the SEFA model presumed here, the next question is very important and pertains to the mapping rule used to squash certain cells of $\boldsymbol{\beta}$ to zero. If a CFA model were being estimated, this questions in figure 8 would not appear because the Howe (1955) theorem would be satisfied by *a priori* restrictions.

Figure 6: (Not) Specifying Values for the Primary Pattern Coefficients at Level 1



	row.names	Factor_1	Factor_2	Factor_3	Factor_4	Factor_5
1	VisualPerception	NA	NA	NA	NA	NA
2	Cubes	NA	NA	NA	NA	NA
3	PaperFormBoard	NA	NA	NA	NA	NA
4	Flags	NA	NA	NA	NA	NA
5	GeneralInformation	NA	NA	NA	NA	NA
6	PargraphComprehension	NA	NA	NA	NA	NA
7	SentenceCompletion	NA	NA	NA	NA	NA
8	WordClassification	NA	NA	NA	NA	NA
9	WordMeaning	NA	NA	NA	NA	NA
10	Addition	NA	NA	NA	NA	NA
11	Code	NA	NA	NA	NA	NA
12	CountingDots	NA	NA	NA	NA	NA
13	StraightCurvedCapitals	NA	NA	NA	NA	NA
14	WordRecognition	NA	NA	NA	NA	NA
15	NumberRecognition	NA	NA	NA	NA	NA
16	FigureRecognition	NA	NA	NA	NA	NA
17	ObjectNumber	NA	NA	NA	NA	NA
18	NumberFigure	NA	NA	NA	NA	NA
19	FigureWord	NA	NA	NA	NA	NA
20	Deduction	NA	NA	NA	NA	NA
21	NumericalPuzzles	NA	NA	NA	NA	NA
22	ProblemReasoning	NA	NA	NA	NA	NA
23	SeriesCompletion	NA	NA	NA	NA	NA
24	ArithmeticProblems	NA	NA	NA	NA	NA
25						

Figure 7: Specifying Bounds on the Primary Pattern Coefficients at Level 1

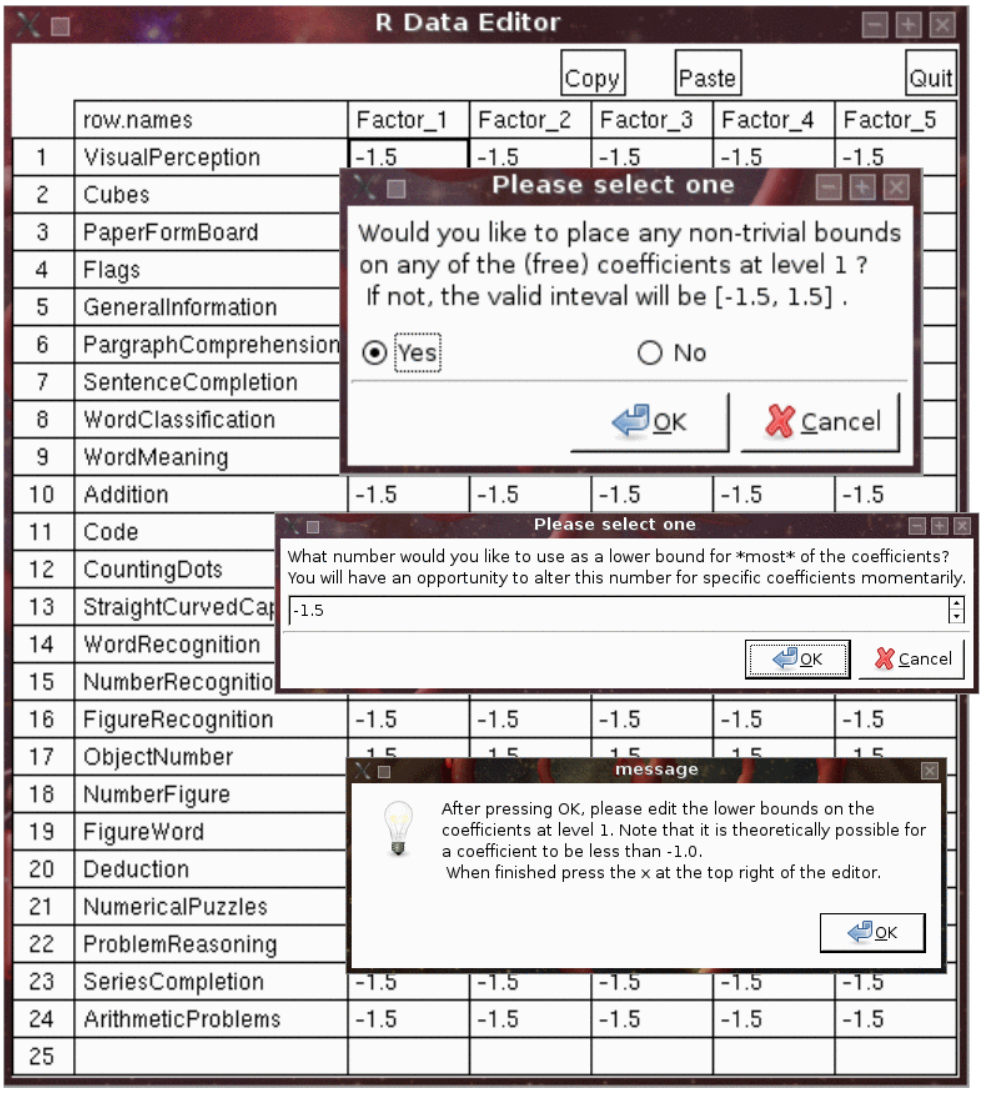
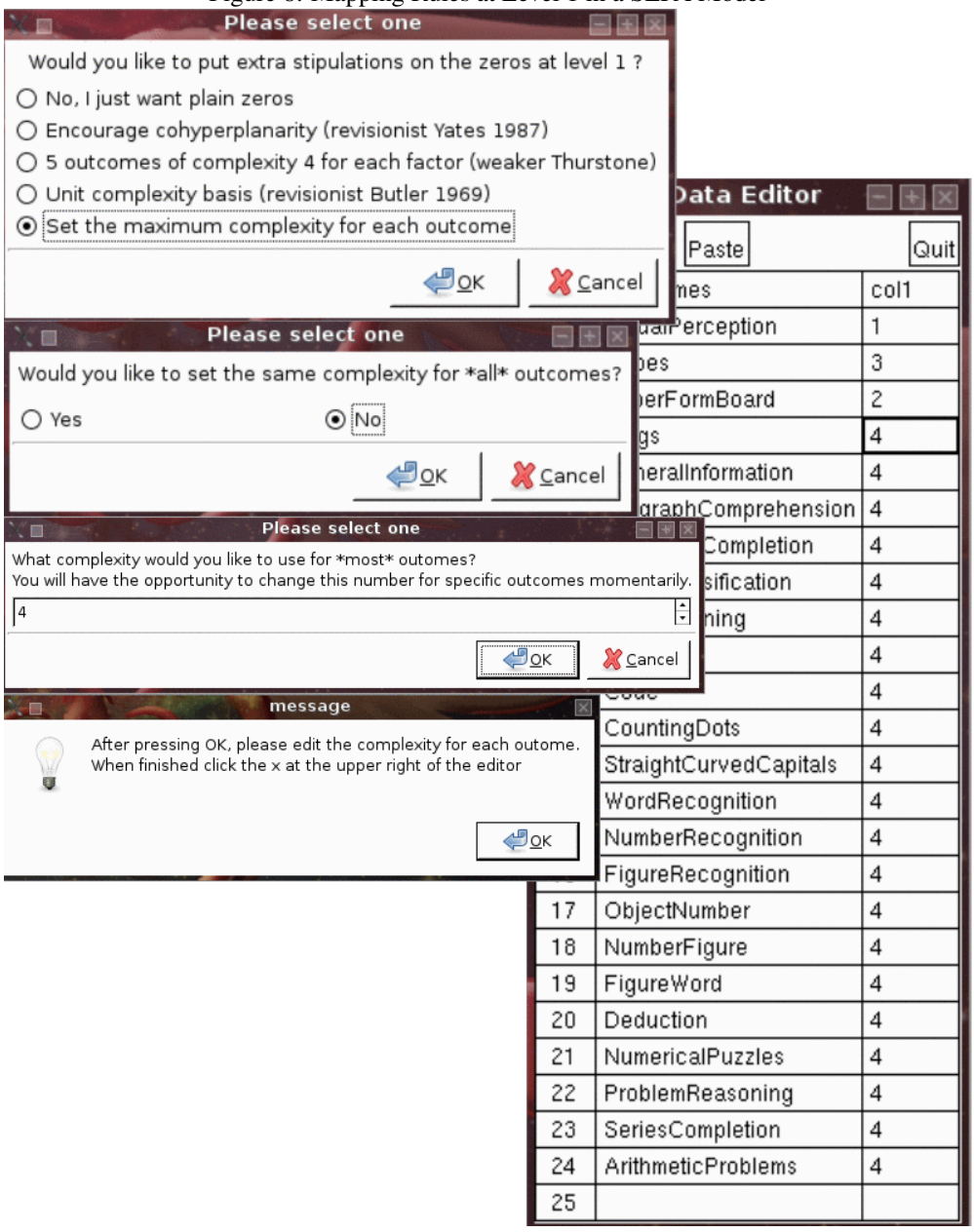


Figure 8: Mapping Rules at Level 1 in a SEFA Model



Let b_p be the number of zeros required for the pth column of $\boldsymbol{\beta}$. Most of the mapping rules at the second level are defined on the reference structure matrix $\left(\check{\boldsymbol{\Upsilon}}\right)$ and the factor contribution matrix $\left(\check{\boldsymbol{\Gamma}}\right)$, rather than the primary pattern matrix ($\check{\boldsymbol{\beta}}$), where the $\check{}$ indicates a prelimary matrix that is constructed by filling the free elments with $\boldsymbol{\theta}$ (in symbols) or `par` (in R syntax). By "smallest", I always mean in absolute value. The mapping rules in the first dialog box in figure 8 are defined as follows:

- "No, I just want plain zeros". This mapping rule simply squashes the b_p smallest elements in the pth column of $\check{\boldsymbol{\Upsilon}}$ to zero, implying that $\boldsymbol{\beta}_p$ has b_p zeros.
- "Encourage cohyperplanrity (revisionist Yates 1987)". This mapping rule squashes $r_2 - 1$ elements in the pth column of $\boldsymbol{\beta}$ using the following piecewise function $\forall q \neq p$:

$$\beta_{jp} \quad = \quad \begin{cases} 0 & \text{if } \check{\Gamma}_{jq} - \check{\Gamma}_{jq} = \max\left\{\check{\boldsymbol{\Gamma}}_q - \check{\boldsymbol{\Gamma}}_p\right\} \\ \check{\beta}_{jp} & \text{otherwise,} \end{cases}$$

  which tends to place the zeros for the pth column of $\boldsymbol{\beta}$ in the rows where the factor contribution of some other column is large.
- "r_1 outcomes of complexity $r_1 - 1$ for each factor (weaker Thurstone)". This mapping rule is similar to the "plain zeros" mapping rule in that r_1 small elements in the pth column of $\check{\boldsymbol{\Upsilon}}$ are squashed to zero, except that there is an additional stipulation that no outcome is of complexity less than $r_1 - 1$ where the complexity of the jth outcome is the number of non-zeros in $\boldsymbol{\Upsilon}_j$. However, any "non-zero" coefficient can be arbitrarily close to zero. Thus, each column of $\boldsymbol{\beta}$ will have r_1 *exact* zeros.
- "Unit complexity basis (revisionist Butler 1969)". This mapping rule uses the dishinguishability weights defined in Yates (1987) to select r_1 outcomes to be of unit complexity. In particular, let

$$\check{w}_p \quad = \quad \frac{1}{\sum \Gamma_{\bullet\bullet}} \sum_{j=1}^{n} \check{\Gamma}_{jp}$$

  where $\sum \Gamma_{\bullet\bullet} = \sum_{j=1}^{n} \sum_{p=1}^{r_1} \check{\Gamma}_{jp}$ is the weight applied the the pth first-order factor such that $\sum_{p=1}^{r_1} w_p = 1$. This equation is similar to equation 118 in Yates (1987, p. 145). Then let

$$\check{c}_j \quad = \quad \frac{\sum_{p=1}^{r_1} \check{w}_p \check{\beta}_{jp}}{\sum_{p=1}^{r_1} \check{\Gamma}_{jp}}$$

  be the directional cosine between $\check{\boldsymbol{\beta}}_j$ and the (weighted) centroid of $\check{\boldsymbol{\beta}}$, which is the same as equation 119 in Yates (1987, p. 145). Finally, let

$$d_j^w \quad = \quad \min\left\{1, \frac{(r_1)^2}{r_1 - 1} \check{c}_j \left(\check{c}_j - 1\right)\right\}$$

  be the distinguishability weight for $\check{\boldsymbol{\beta}}_j$, which is the same as equation $78a$ in Yates (1987, p. 112). Then, $\forall p$ choose a outcome to be of unit complexity among outcomes where $\check{\Gamma}_{jp} = \max\left\{\check{\boldsymbol{\Gamma}}_j\right\}$ such that d_j^w is maximal. For such outcomes, squash $\boldsymbol{\beta}_j$ to zero $\forall q \neq p$. At the end of these steps, there will be r_1 outcomes that are collinear with the first-order primary axex and there will be $r_1 - 1$ zeros in each column of $\boldsymbol{\beta}$.
- "Set the maximum complexity for each outcome." This mapping rule is similar to the "plain zeros" mapping rule, except that it is applied to each *row* of $\boldsymbol{\beta}$ rather than the columns. In general, $\forall j$, the user can choose a number between 1 and r_1 for the maximum number of non-zero coefficients in $\boldsymbol{\beta}_j$. This mapping rule is used as an example in figure 8. It is possible to choose the same complexity for all rows of $\boldsymbol{\beta}$, in which case a simpler dialog box is shown. Note that if the complexity of all rows of $\boldsymbol{\beta}$ is $r_1 - 1$, this restriction characterizes

Thurstone's (1935) definition of simple structure, and if the complexity of all rows is 1, this restriction implies a perfect cluster configuration. In figure 8, I have chosen to (arbitrarily) specify the complexity of the rows of $\boldsymbol{\beta}$ on a row-by-row basis, which gives more flexibility to place restrictions on the model in accordance with prior theory.

In some cases, b_p may be greater than the number of zeros created by whichever mapping rule is chosen. In that case, the remaining zeros are filled in by squashing the b_p smallest elements in the pth column of the *recalculated* reference structure matrix after the mapping rule has been applied and some reference structure correlations are zero. Thus, it is necessary to specify b_p. By default, the user is asked whether to set $b_p = r_1 \, \forall p$, which can be accepted by answering "Yes" to the question in figure 8.

2.3 Lexical Criteria

Finally, the user is asked which lexical criteria should be included in the optimization. Regardless of what options are chosen here, there will always be criteria that impose necessary constraints on the model (positive definite correlation matrices, positive specific variances, etc.) and constraints that the column sums of $\boldsymbol{\Delta}$ (if a second-order model is estimated) and $\boldsymbol{\beta}$ are positive, which improves the computational efficiency of the algorithm. Also, regardless of what options are chosen here, the ultimate criterion will be appended to the end of the criteria that define constraints. In this case, the model is to be estimated by maximum likelihood, so the dialog box reminds the user of this. If a second-order model is not estimated, the choices that mention with "level 2" in figure 9 would not appear.

Peña and Rodriguez (2003) defines a determinant of a correlation matrix among a set of variables raised to the power of the reciprocal of the number of variables as the "effective variance" of those variables. The essence of the effective variance is that it can be meaningfully compared to a different set of variables when the sets of variables differ in size. The generalized variance is more well-known statistic and is simply the determinant of the correlation matrix among a set of variables. The generalized variance can be interpreted as the hypervolume of a set of unit-length vectors and the effective variance can be interpreted as "the length of the side of the hypercube whose volume is equal to the determinant (Peña and Rodriguez 2003, p. 363)." Both of these concepts are used to define the criteria / constraints in figure 9:

- "Reference factors at level 2 must have more effective variance than do the primary factors at level 1". This criterion is formally defined as

$$\begin{cases} 1 & \text{if } |\boldsymbol{\Phi}|^{\frac{1}{r_1}} \leq \left|\mathbf{Z}\boldsymbol{\Xi}^{-1}\mathbf{Z}\right|^{\frac{1}{r_2}} \\ |\boldsymbol{\Phi}|^{\frac{1}{r_1}} - \left|\mathbf{Z}\boldsymbol{\Xi}^{-1}\mathbf{Z}\right|^{\frac{1}{r_2}} & \text{otherwise,} \end{cases}$$

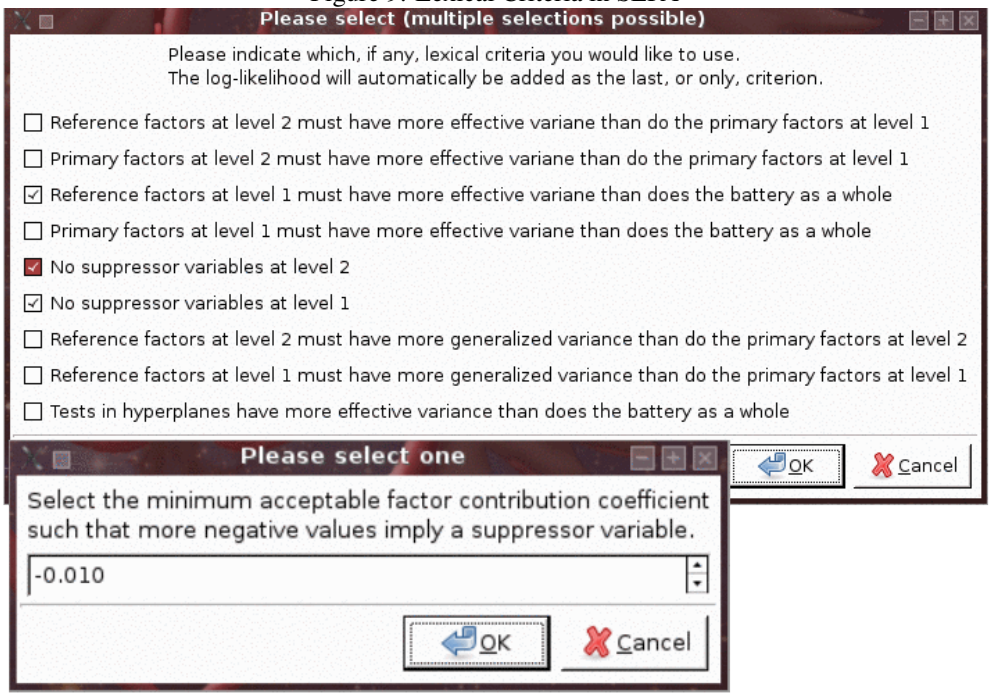
  and loosely requires that the reference factors at level 2 be "less correlated per variable" than are the primary factors at level 1. This criterion tends to push the second-order axes toward the edge of the configuration of first-order factors.

- "Primary factors at level 2 must have more effective variance than do the primary factors at level 1". This criterion is formally defined as

$$\begin{cases} 1 & \text{if } |\boldsymbol{\Phi}|^{\frac{1}{r_1}} \leq |\boldsymbol{\Xi}|^{\frac{1}{r_2}} \\ |\boldsymbol{\Phi}|^{\frac{1}{r_1}} - |\boldsymbol{\Xi}|^{\frac{1}{r_2}} & \text{otherwise,} \end{cases}$$

  and is similar to the criterion above. In general, it seems that imposing this restriction on the second-order primary factors is *stronger* than imposing it on the second-order reference factors and pushes the second-order axes farther toward the edge of the configuration of first order factors.

Figure 9: Lexical Criteria in SEFA



- "Reference factors at level 1 must have more effective variance than does the battery as a whole". This criterion is formally defined as

$$\begin{cases} 1 & \text{if } |\mathbf{\Psi}|^{\frac{1}{r_1}} \leq |\mathbf{C}|^{\frac{1}{n}} \\ |\mathbf{\Psi}|^{\frac{1}{r_1}} - |\mathbf{C}|^{\frac{1}{n}} & \text{otherwise,} \end{cases}$$

  and is similar to the two criteria above except that the comparison is relative to $\mathbf{C}$. In this case, the criterion loosely implies that the reference factors at level 1 be "less correlated per variable" than is the test configuration in common factor space and encourages the first-order axes to cut through the edges of the test configuration.

- "Primary factors at level 1 must have more effective variance than does the battery as a whole". This criterion is formally defined as

$$\begin{cases} 1 & \text{if } |\mathbf{\Phi}|^{\frac{1}{r_1}} \leq |\mathbf{C}|^{\frac{1}{n}} \\ |\Phi|^{\frac{1}{r_1}} - |\mathbf{C}|^{\frac{1}{n}} & \text{otherwise,} \end{cases}$$

  and is similar to the previous criterion above except that the correlation among the primary factors at issue. In general, it seems that imposing this restriction on the primary factors is *stronger* than imposing it on the reference factors and tends to push the first-order axes farther out into the edges of the test configuration.

- "No suppressor variables at level 2" Suppressor variables are defined as variables with negative factor contributions. This criterion is advocated, albeit somewhat indirectly, in Yates (1987, p. 119) with respect to level 1, although I can see no reason why it would be any less applicable to level 2. Prohibiting suppressor variables strikes me as a reasonable assumption to make in most cases. The criterion is formally defined as

$$\frac{1}{r_1 r_2} \sum_{j=1}^{r_1} \sum_{p=1}^{r_2} \mathbb{I}\left\{\eth_{jp} \geq \underline{\eth}\right\} \quad \text{where} \quad \mathbb{I}\left\{\eth_{jp} \geq \underline{\eth}\right\} = \begin{cases} 1 & \text{if } \eth_{jp} \geq \underline{\eth} \\ 0 & \text{otherwise,} \end{cases}$$

and $\underline{\supseteq}$ is a threshold for the minimum acceptable factor contribution that is specified by the user, as shown at the bottom of figure 9.

- "No suppressor variables at level 1". This criterion is conceptually similar to the one above and is formally defined at level 1 as

$$\frac{1}{nr}\sum_{j=1}^{n}\sum_{p=1}^{r}\mathbb{I}\left\{\Gamma_{jp}\geq\underline{\Gamma}\right\}\quad\text{where}\quad\mathbb{I}\left\{\Gamma_{pj}\geq\underline{\Gamma}\right\}=\begin{cases}1 & \text{if } \Gamma_{pj}\geq\underline{\Gamma}\\ 0 & \text{otherwise,}\end{cases}$$

  and $\underline{\Gamma}$ is a threshold for the minimum acceptable factor contribution that is specified by the user as shown at the bottom of figure 9.

- "Reference factors at level 2 must have more generalized variance than do primary factors at level 2". This criterion is extrapolated from Yates (1987, p. 27), although it is not asserted with much formality in the text, is not specifically advocated for level 2. Regardless, the criterion could be operationalized as

$$\begin{cases}1 & \text{if } |\boldsymbol{\Xi}|\leq\left|\mathbf{Z}\boldsymbol{\Xi}^{-1}\mathbf{Z}\right|\\ |\boldsymbol{\Xi}|-\left|\mathbf{Z}\boldsymbol{\Xi}^{-1}\mathbf{Z}\right| & \text{otherwise,}\end{cases}$$

  but has a faster and substantively equivalent operationalization in the code. However, when $r_2=2$, $|\boldsymbol{\Xi}|=\left|\mathbf{Z}\boldsymbol{\Xi}^{-1}\mathbf{Z}\right|$ by necessity, so this criterion can bind only in the rare case where $r_2\geq 3$.

- "Reference factors at level 1 must have more generalized variance than do primary factors at level 1". This criterion is conceptually the same as the previous one and could be operationalized as

$$\begin{cases}1 & \text{if } |\boldsymbol{\Phi}|\leq|\boldsymbol{\Psi}|\\ |\boldsymbol{\Phi}|-|\boldsymbol{\Psi}| & \text{otherwise,}\end{cases}$$

  but has a faster and substantively equivalent operationalization in the code. When $r_1=2$, $|\boldsymbol{\Phi}|=|\boldsymbol{\Psi}|$ by necessity, so this criterion can bind in the case when $r_1\geq 3$.

- "Tests in hyperplanes have more effective variance than does the battery as a whole". This criterion is only defined for level 1, although I suppose it could be defined for level 2 at some point in the future. It seems consistent with the spirit of Yates (1987) but is not explicitly advocated anywhere. This criterion is formally defined as

$$\frac{1}{r_1}\sum_{p=1}^{r_1}\mathbb{I}\left\{\left|\overset{p}{\boldsymbol{\beta}}\boldsymbol{\Phi}\overset{p}{\boldsymbol{\beta}}'+\overset{p}{\boldsymbol{\Theta}}^2\right|^{\frac{1}{b}}\geq|\mathbf{C}|^{\frac{1}{n}}\right\}\quad\text{where}\quad\mathbb{I}\left\{\left|\overset{p}{\boldsymbol{\beta}}\boldsymbol{\Phi}\overset{p}{\boldsymbol{\beta}}'+\overset{p}{\boldsymbol{\Theta}}^2\right|^{\frac{1}{b}}\geq|\mathbf{C}|^{\frac{1}{n}}\right\}=\begin{cases}1 & \text{if } \left|\overset{p}{\boldsymbol{\beta}}\boldsymbol{\Phi}\overset{p}{\boldsymbol{\beta}}'+\overset{p}{\boldsymbol{\Theta}}^2\right|^{\frac{1}{b}}\geq|\mathbf{C}|^{\frac{1}{n}}\\ 0 & \text{otherwise,}\end{cases}$$

  $\overset{p}{\boldsymbol{\beta}}$ is the $b_p\times r_1$ submatrix of $\boldsymbol{\beta}$ with exact zeros in the pth column and $\overset{p}{\boldsymbol{\Theta}}^2$ is the $b_p\times b_p$ diagonal (but not necessarily contiguous) submatrix of $\boldsymbol{\Theta}^2$ with the unique variances of the outcomes in the pth hyperplane. Thus, this criterion essentially requires that the tests in the pth hyperplane be "less correlated per variable" than is the battery as a whole (in common factor space) for all r_1 hyperplanes at level 1.

Note that it is possible to select more than one of these criteria as constraints in the lexical optimization process.

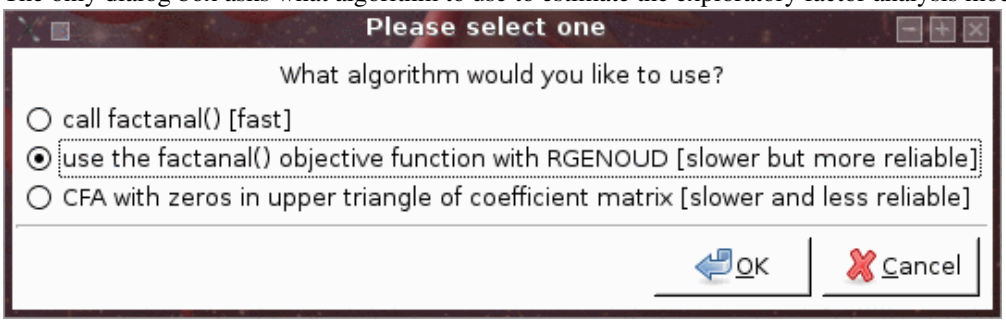
3 Exploratory Factor Analysis via `Factnal()` and `Rotate()`

This strategy of EFA is explained in more detail in Goodrich (2008b). To start this sequence, I typed the command
`mental.tests.efa <- Factanal(covmat = Harman74.cor, factors = 5, model = "EFA")`

3.1 Factor Extraction

I use $\mathbf{\Lambda}$ to indicate the preliminary pattern matrix to estimate when the preliminary factors are orthogonal.

The only dialog box asks what algorithm to use to estimate the exploratory factor analysis model:



The first option simply calls `factanal()`, which is fine for many problems and estimates the EFA model with the restrictions that $\mathbf{\Lambda}'\mathbf{\Theta}^2\mathbf{\Lambda}$ is diagonal and $\mathbf{\Phi} = \mathbf{I}$. The second algorithm imposes these same restrictions but uses `genoud()` to optimize with a genetic algorithm. This approach is much slower but slightly more likely to reach the global optimum in complicated problems The third algorithm estimates the model under restrictions that $\mathbf{\Lambda} = \begin{bmatrix}\mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_2\end{bmatrix}$ where $\mathbf{\Lambda}_1$ is a $r_1 \times r_1$ matrix with zeros above the diagonal and $\mathbf{\Phi} = \mathbf{I}$.

In principle, all three algorithms should produce the same estimate of $\mathbf{\Theta}^2$ and be within a rotation of each other. In practice, the third algorithm has difficulty meeting these standards.

3.2 Factor Transformation

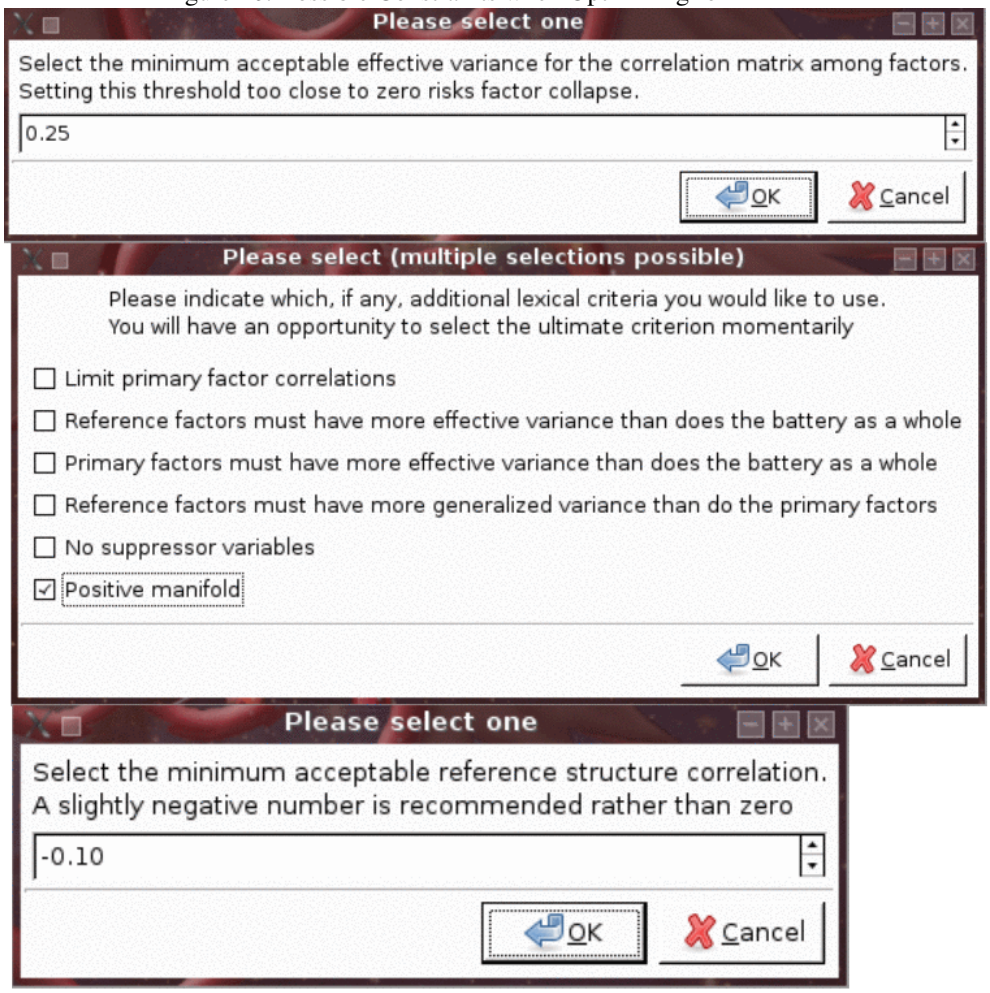
Let $\mathbf{\Phi} = \mathbf{T}'\mathbf{T}$ where $\mathbf{T}$ is a $r_1 \times r_1$ transformation matrix with unit-length columns. The factor transformation step chooses $\mathbf{T}$ so so that some (lexical) objective function is *minimized*. I initiated this sequence with
`mental.tests.rotated <- Rotate(mental.tests.efa)`

3.2.1 Constraints

Peña and Rodriguez (2003) defines a determinant of a correlation matrix among a set of variables raised to the power of the reciprocal of the number of variables as the "effective variance" of those variables. The essence of the effective variance is that it can be meaningfully compared to a different set of variables when the sets of variables differ in size. The generalized variance is more well-known statistic and is simply the determinant of the correlation matrix among a set of variables. The generalized variance can be interpreted as the hypervolume of a set of unit-length vectors and the effective variance can be interpreted as "the length of the side of the hypercube whose volume is equal to the determinant (Peña and Rodriguez 2003, p. 363)." Both of these concepts are used to define the criteria / constraints in figure 10 to be used while seeking the optimal $\mathbf{T}$.

The first criterion is required and prevents factor collapse. The effective variance of the primary factors is defined as $|\mathbf{\Phi}|^{\frac{1}{r_1}}$. If $|\mathbf{\Phi}|^{\frac{1}{r_1}} = 0$, then $\mathbf{\Phi}$ and $\mathbf{T}$ are singular, a circumstance which is known as "factor collapse". By specifying

Figure 10: Possible Constraints when Optimizing for $\mathbf{T}$ in EFA



a non-zero threshold that the effective variance must exceed, factor collapse can thus be prevented. In this case, the threshold is 0.25 and the criterion is formally defined as

$$\begin{cases} -1 & \text{if } |\mathbf{\Phi}|^{\frac{1}{r_1}} \geq 0.25 \\ -\left(|\mathbf{\Phi}|^{\frac{1}{r_1}}\right) & \text{otherwise.} \end{cases}$$

The remaining constraints are optional and are enumerated in the next dialog box in figure 10. These criteria are defined as follows:

- "Limit primary factor correlations". This option allows the user to set valid intervals for pairwise corrleations between primary factors, i.e. the off-diagonals of $\mathbf{\Phi} = \mathbf{T}'\mathbf{T}$. If this box is checked, dialog boxes that are similar to the previous one will pop up asking for minimum and maximum acceptable correlations. The associated criterion equals -1 if all off-diagonals are within their valid intervals and equal to the distance between the most abberrant correlation and the valid interval otherwise.

- "Reference factors must have more effective variance than does the battery as a whole". This criterion is formally defined as

  $$\begin{cases} 1 & \text{if } |\mathbf{\Psi}|^{\frac{1}{r_1}} \leq |\mathbf{C}|^{\frac{1}{n}} \\ |\mathbf{\Psi}|^{\frac{1}{r_1}} - |\mathbf{C}|^{\frac{1}{n}} & \text{otherwise.} \end{cases}$$

  This criterion loosely implies that the reference factors must be "less correlated per variable" than is the test configuration in common factor space and encourages the first-order axes to cut through the edges of the test configuration.

- "Primary factors must have more effective variance than does the battery as a whole". This criterion is formally defined as

  $$\begin{cases} 1 & \text{if } |\mathbf{\Phi}|^{\frac{1}{r_1}} \leq |\mathbf{C}|^{\frac{1}{n}} \\ |\Phi|^{\frac{1}{r_1}} - |\mathbf{C}|^{\frac{1}{n}} & \text{otherwise,} \end{cases}$$

  and is similar to the previous criterion above except that the correlation among the primary factors at issue. In general, it seems that imposing this restriction on the primary factors is *stronger* than imposing it on the reference factors and tends to push the first-order axes farther out into the edges of the test configuration.

- "Primary factors must have less generalized variance than do reference factors". This criterion is extrapolated from Yates (1987, p. 27), although it is not asserted with much formality in the text. Regardless, the criterion could be operationalized as

  $$\begin{cases} -1 & \text{if } |\mathbf{\Phi}| \leq |\mathbf{\Psi}| \\ |\mathbf{\Psi}| - |\mathbf{\Phi}| & \text{otherwise,} \end{cases}$$

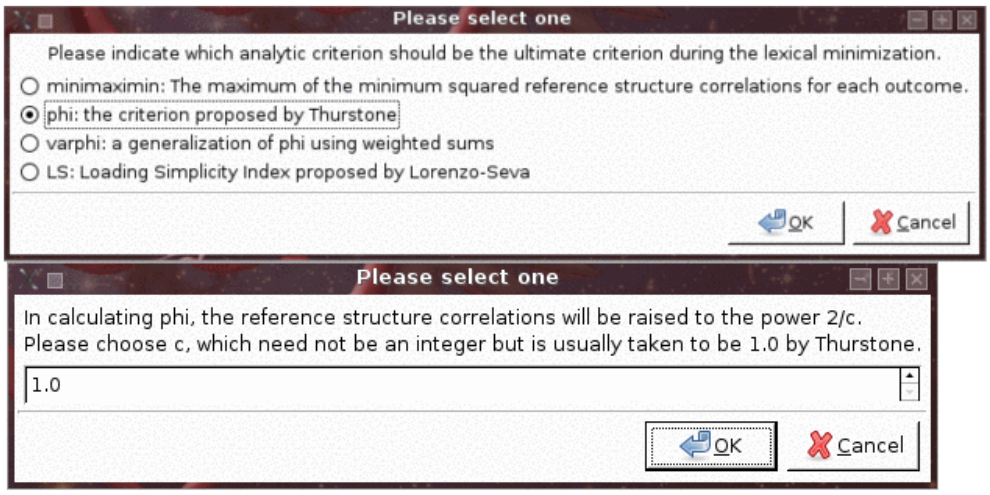
  but has a faster and substantively equivalent operationalization in the code. When $r_1 = 2$, $|\mathbf{\Phi}| = |\mathbf{\Psi}|$ by necessity, so this criterion can bind only when $r_1 \geq 3$.

- "No suppressor variables". Suppressor variables are defined as variables with negative factor contributions. This criterion is advocated, albeit somewhat indirectly, in Yates (1987, p. 119). Prohibiting suppressor variables strikes me as a reasonable assumption to make in most cases. The criterion is formally defined as

  $$-\frac{1}{nr}\sum_{j=1}^{n}\sum_{p=1}^{r}\mathbb{I}\left\{\Gamma_{jp} \geq \underline{\Gamma}\right\} \quad \text{where} \quad \mathbb{I}\left\{\Gamma_{jp} \geq \underline{\Gamma}\right\} = \begin{cases} 1 & \text{if } \Gamma_{jp} \geq \underline{\Gamma} \\ 0 & \text{otherwise,} \end{cases}$$

  and $\underline{\Gamma}$ is a threshold for the minimum acceptable factor contribution that is specified by the user. If this criterion is checked, a pop-up will ask for the value of $\underline{\Gamma}$. See the very end of section 2.2.3 for a picture and explanation of this dialog box, which is virtually identical to the next one.

Figure 11: Ultimate Criteria to be Used when Optimizing for $\mathbf{T}$ in EFA



- "Positive manifold". This criterion is similar to (and a bit more restrictive than) the criterion that prohibits suppressor variables. The criterion is formally defined as

$$-\frac{1}{nr}\sum_{j=1}^{n}\sum_{p=1}^{r}\mathbb{I}\left\{\Upsilon_{jp}\geq\underline{\Upsilon}\right\} \quad \text{where} \quad \mathbb{I}\left\{\Upsilon_{jp}\geq\underline{\Upsilon}\right\}=\begin{cases}1 & \text{if } \Upsilon_{jp}\geq\underline{\Upsilon}\\ 0 & \text{otherwise,}\end{cases}$$

  where $\underline{\Upsilon}$ is a threshold for the minimum acceptable reference structure correlation that is specified by the user. If this criterion is checked, the dialog box at the bottom of figure 10will ask for the value of $\underline{\Upsilon}$.

3.2.2 The Ultimate Criterion

The final step is to choose the ultimate criterion for lexical optimization, as shown in the dialog box in figure 11. These criteria are defined as follows:

- minimaximin: This criterion attempts to satisfy Thurstone's definition of simple structure by choosing $\mathbf{T}$ so that all outcomes have at least one near-zero reference structure correlation. The criterion is formally defined as

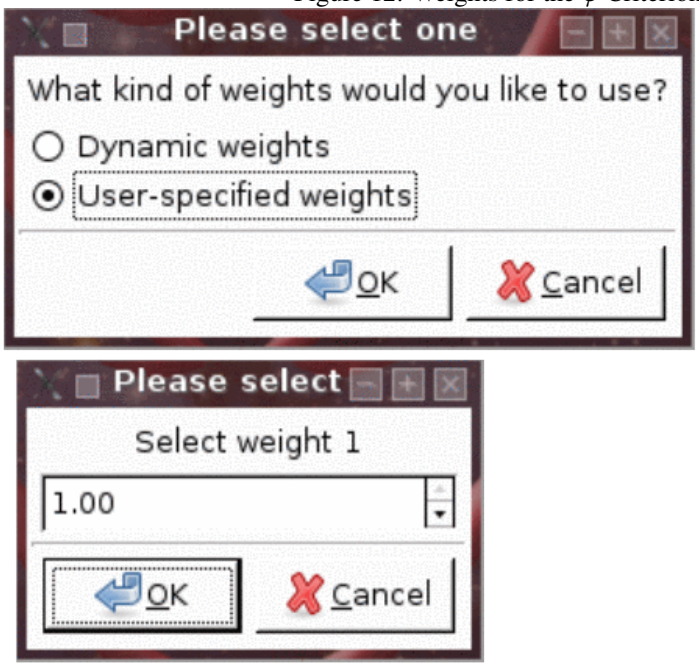
$$\ln\left(\max\left\{\min\left\{\Upsilon_1^2\right\},\min\left\{\Upsilon_2^2\right\},\ldots,\min\left\{\Upsilon_n^2\right\}\right\}\right).$$

  Thus, the (log of the) maximum of the minimums of the squared rows of $\boldsymbol{\Upsilon}$ is minimized. Taking the natural logarithm just makes the optimization work a bit better near zero.

- phi: This criterion was proposed by Thurstone (1935) to numerically characterize simple structure when it reaches its theoretical minimum of (log) zero. The criterion is formally defined as

$$\ln\left(\sum_{j=1}^{n}\exp\left(\tfrac{1}{c}\sum_{p=1}^{r_1}\ln\left(\Upsilon_{jp}^2\right)\right)\right),$$

Figure 12: Weights for the φ Criterion



which is equal to logarithm of ϕ as defined by Thurstone (1935) when $c = 1$. Thurstone suggested that choosing c to be greater than 1.0 could yield better results by lessening the pressure to get one near-zero into each row of $\boldsymbol{\Upsilon}$ and increasing the pressure to get more than one near-zero into some rows of $\boldsymbol{\Upsilon}$. If Thurstone's criterion is used, the dialog box at the bottom of figure 11will appear, prompting the user for c.

- varphi: This criterion is a generalization of Thurstone's ϕ. Let $\overrightarrow{\boldsymbol{\Upsilon}}$ be a matrix of reference structure correlations where each row is independently sorted in increasing magnitude, and let $\underset{-[1:p]}{\overrightarrow{\phi}}$ be Thurstone's criterion (with $c = 1$) calculated on $\overrightarrow{\boldsymbol{\Upsilon}}$, excluding the first through the pth column of $\overrightarrow{\boldsymbol{\Upsilon}}$. Then define $\varphi = \phi + \sum_{p=1}^{r-2} w_p \underset{-[1:p]}{\overrightarrow{\phi}}$, where $w_p \in [0, 1]$ is the weight the analyst specifies for $\underset{-[1:p]}{\overrightarrow{\phi}}$ relative to a unit weight placed on ϕ. At one extreme, if $w_p = 0 \, \forall p$, then $\varphi = \phi$. At the other extreme, if $w_p = 1 \, \forall p$, then the analyst favors a perfect cluster configuration, where three is only one non-zero coefficient per test. An "objective" alternative to specifying weights is to specify that the weights are a function of $\overrightarrow{\boldsymbol{\Upsilon}}$, such as $w_p = \max\left\{0, 1 - \max\left\{\overrightarrow{\Upsilon}^2_{(p+1)}\right\}\right\}$ to gradually reduce the weight as p increases. If this criterion is chosen, the dialog boxes in figure 12 will appear, prompting the user to specify whether dynamic weights or static weights should be used. If user-specified weights are chosen, the user is then prompted to specify each w_p .

- LS: This criterion is called the Loading Simplicy Index in Lorenzo-Seva (2003). The derivation of it is not complicated but involves a lot of notation not previously introduced here, so the reader is refered to Lorenzo-Seva (2003) for details. Simply put, it reaches its optimum of (negative) 1.0 when $\boldsymbol{\Upsilon}$ exhibits a perfect cluster configuration. Although Lorenzo-Seva (2003) defines the Loading Simplicity Index in terms of the primary pattern matrix, it is invariant to the column-scale of the primary pattern matrix and thus can arbitrarily be defined for the reference structure matrix as well.

4 References

Butler, J.M. 1969. "Simple Structure Reconsidered: Distinguishability and Invariance in Factor Analysis." *Multivariate Behavioral Research* 4(1):5–28.

Goodrich, B. 2008a. "SEFA*i*R So Far." Unpublished manuscript available from the FAiR wiki `http://wiki.r-project.org/rwiki/doku.php?id=packages:cran:fair`

Goodrich, B. 2008b. "Analytic Transformation of Factors in FA*i*R." Unwritten manuscript. Email me about it.

Howe, W.G. 1955. Some Contributions to Factor Analysis. Technical report ORNL-1919, Oak Ridge National Lab., Tenn.

Lorenzo-Seva, U. 2003. "A factor simplicity index." *Psychometrika* 68(1):49–60.

Peña, D. and J. Rodriguez. 2003. "Descriptive measures of multivariate scatter and linear dependence." *Journal of Multivariate Analysis* 85(2):361–374.

Thurstone, L.L. 1935. *The vectors of mind: multiple-factor analysis for the isolation of primary traits*. University of Chicago Press.

Yates, A. 1987. *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. State University of New York Press.