Studies on the effect of violations of local independence on scale in Rasch models: The Dichotomous Rasch model

Ida Marais

and

David Andrich

Murdoch University, Western Australia

Mailing address

David Andrich

Murdoch University

Murdoch 6150

Western Australia

Acknowledgements

The work for the Report was supported in part by an Australian Research Council grant with the Australian National Ministerial Council on Employment, Education, Training and Youth Affairs (MCEETYA) Performance Measurement and Reporting Task Force, UNESCO's International Institute for Educational Planning (IEP), and the Australian Council for Educational Research (ACER) as Industry Partners*.

*Report No. 1 ARC Linkage Grant LP0454080: Maintaining Invariant Scales in State, National and International Level Assessments. D Andrich and G Luo Chief Investigators, Murdoch University

Studies on the effect of violations of local independence on scale in Rasch models: The Dichotomous Rasch model

1. Introduction

The unidimensional Rasch model

$$\Pr\{x_{ni}\} = [\exp(x_{ni}(\beta_n - \delta_i) - \sum_{k=1}^{x}\tau_{ki})] / \sum_{x=0}^{m_i}[\exp(x_{ni}(\beta_n - \delta_i) - \sum_{k=1}^{x}\tau_{ki})]), \qquad (1)$$

implies independence of responses in the sense that

$$\Pr\{((x_{ni}))\} = \prod_n \prod_i \Pr\{x_{ni}\} \qquad (2)$$

where $x_{ni} \in \{0, 1, 2 \ldots m_i\}$ is the integer response variable for person *n* with ability β_n responding to item with difficulty δ_i, and thresholds between the graded responses $\tau_{1i}, \tau_{2i,\ldots}\ \tau_{mi}$, and where m_i is the maximum score of item *i*. This implies a single dimension with *β, δ* and *τ* located on the same scale.

No data fit any model perfectly and there are concerns about violations of these two conditions (independence and unidimensionality) in the Rasch model of Eq. (1). This paper develops a simulation rationale for two types of dependence that we have called *trait dependence* and *response dependence*. Data is simulated with varying levels of dependence and then analysed according to the model to show typical effects in the analyses from these specific violations of the model. The effects of interest are effects on the unit of scale, as manifested in the range of the person locations and the standard deviation of person locations, and effects on typical diagnostics of statistical indices of misfit. Trait dependence and unidimensionality are reflections of each other in that violation of one implies a particular violation of the other. The simulation rationale for trait dependence shows this relationship.

2. Simulation algorithms for dependency

A simulation algorithm for trait dependency

Assessments often involve more than one component – for example, a multiple choice items component, a written component, and a performance component. It would be expected that the components of an assessment will have a reasonably high positive correlation. They will not be perfectly correlated, because, if they were, there would be no need to identify the distinct components – there would be just one component. Each assessment component has a set of items in which the latent traits of the components are correlated amongst each other to varying degrees. Some components may have only one item.

2.1 Simulation of two assessment components

Consider the case of an assessment with two components. Let B_1 and B_2 be the two latent variables which are assessed by two sets of items, called components above. The latent variable B_1 is involved in responding to component 1 and the latent ability B_2 is involved in responding to component 2. B_1 and B_2 are correlated to varying degrees, sometimes approaching 1. Let β, β_1 and β_2 be three other variables which are not correlated with each other. Let β be the common component of B_1 and B_2, which is the source of the correlation between them. Let β_1 and β_2 reflect the unique aspects of each of the components of items. Let the distributions of β, β_1 and β_2 be identical, normally distributed with mean 0 and standard deviation 1.

To construct a value for B_1 and B_2 for each person, the first step is to simulate three independent standard normal random deviates β, β_1 and β_2.

Then define $B_1 = a_1 + \frac{b_1}{\sqrt{1+c^2}} A_1$ and $B_2 = a_2 + \frac{b_2}{\sqrt{1+c^2}} A_2$ (3)

where $A_1 = \beta + c\,\beta_1$ and $A_2 = \beta + c\,\beta_2$

The source of the correlation between B_1 and B_2 is the common latent variable β; the source of the correlation not being 1.0 is the presence of β_1 and β_2 with c>0. With c>0 independence is violated because the correlation among item responses within a component is greater than the correlation among items from different components (This is shown in the appendix under *Special: correlation within a component*). In addition to violating local independence it also violates unidimensionality.

It can be shown that B_1 and B_2 have the respective means a_1 and a_2, respective standard deviations b_1 and b_2, and a correlation $r_{12}= \dfrac{1}{1+c^2}$ and $c^2 = \dfrac{1-r_{12}}{r_{12}}$. (This is shown in the appendix). Using these definitions and relationships we can generate latent variables B_1 and B_2 which have any correlation, mean and standard deviation.

2.2 Extending the simulation algorithm: more than two components

2.2.1 More than two components and common correlation

If we require, say, three components of items, define B_1 and B_2 as above, and define a third variable B_3 so that

$$B_3 = a_3 + \frac{b_3}{\sqrt{1+c^2}} A_3 \text{ where } A_3 = \beta + c\,\beta_3$$

If we require the *same correlation* between components, for example all components to be correlated at r = 0.6 then c = $\sqrt{\dfrac{1-r}{r}}$ = 0.82. Note that since r is the same between all components the same constant c is used to define B_1, B_2 and B_3.

2.2.2 More than two components and different correlations

Now consider the case of three components of items with *different correlations* among the components: r_{12}, r_{13} and r_{23}. Since $r_{12} \neq r_{13} \neq r_{23}$ the constant values used to define B_1, B_2 and B_3 will be different as well and defined as c_1, c_2 and c_3.

Let $B_1 = a_1 + \frac{b_1}{\sqrt{1+c_1^2}} A_1$ where $A_1 = \beta + c_1 \beta_1$

and $B_2 = a_2 + \frac{b_2}{\sqrt{1+c_2^2}} A_2$ where $A_2 = \beta + c_2 \beta_2$, etc.

To simplify and generalise, define $E_1 = \frac{1}{\sqrt{1+c_1^2}}$, $E_2 = \frac{1}{\sqrt{1+c_2^2}}$, and $E_3 = \frac{1}{\sqrt{1+c_3^2}}$.

Then it can be shown that $r_{12} = E_1E_2$, $r_{13} = E_1E_3$ and $r_{23} = E_2E_3$.

In the case of three components of items the three correlations are independent of each other. However, consider the case of four components that are all correlated differently with each other. This results in six correlations (r_{12}, r_{13}, r_{14}, r_{23}, r_{24} and r_{34}) that then define c_1, c_2, c_3 and c_4. These six correlations can not be independent of each other and in the correlation matrix there will be relationships among the correlations. The correlation matrix is defined as in Table 1.

Table 1 Correlation matrix for four components

Component	**1**	**2**	**3**	**4**
1	$r_{11} = 1$	$r_{12} = E_1E_2$	$r_{13} = E_1E_3$	$r_{14} = E_1E_4$
2	$r_{21} = E_2E_1$	$r_{22} = 1$	$r_{23} = E_2E_3$	$r_{24} = E_2E_4$
3	$r_{31} = E_3E_1$	$r_{32} = E_3E_2$	$r_{33} = 1$	$r_{34} = E_3E_4$
4	$r_{41} = E_4E_1$	$r_{42} = E_4E_2$	$r_{43} = E_4E_3$	$r_{44} = 1$

Note that given all the four values of E (E_1, E_2, E_3 and E_4) in one row or one column of the matrix the entire six correlation coefficients can be calculated. Thus six correlations are generated from four latent variables, indicating that the six correlations cannot be entirely independent.

Using this simulation rationale we can generate data sets with components of items; the latent variables underlying these components can be correlated according to a given correlation. In the case of more than two components the traits underlying the components can be correlated equally or the traits can have different correlations with

each other, though when the number of components is greater than three, the correlations are not totally independent.

A simulation algorithm for response dependency

Two (or more) items can be dependent on each other in a different way from that generated above, which we have called *response dependence*. An item is dependent on another item in this way when a person's correct response on that item increases the probability of a correct response on the dependent item. Conversely, an incorrect response on that item decreases the probability of a correct response on the dependent item. How much the probability increases or decreases can be determined by a constant c. In the simulation algorithm for trait dependence the person ability β is changed through the constant c. In simulating response dependence the algorithm affects the item difficulty δ through the constant c being added to or subtracted from the difficulty.

Therefore, if a person's response to an item was 0 ($X_i = 0$) then the dependent item's difficulty is changed to $\delta + c$. Hence the dependent item has been made more difficult which increases the probability of an incorrect response to that item for this person. Conversely, if a person's response to an item was 1 ($X_i = 1$) then the dependent item's difficulty is changed to $\delta - c$. Hence the dependent item has been made easier which increases the probability of a correct response to that item for this person. Equation 5 formalises this for item j dependent on item i:

$$\Pr\{X_{nj} = 1 \mid X_{ni} = 1\} = \frac{e^{(\beta_n - (\delta_j - c))}}{1 + e^{(\beta_n - (\delta_j - c))}} \qquad (5)$$

and then $\Pr\{X_{nj} = 1 \mid X_{ni} = 0\} = \dfrac{e^{(\beta_n - (\delta_j + c))}}{1 + e^{(\beta_n - (\delta_j + c))}}$

3. Simulations

Using the simulation rationale described in section 2 data sets were generated with trait dependence and response dependence of varying levels. A data set was also

generated for a baseline condition with no dependence. No condition had both response and trait dependence at the same time. In all the conditions data was generated for 1000 people. The choice of item difficulties and the choice of the distribution of person abilities were meant to make the targeting ideal and the distribution of items symmetrical. Person abilities were normally distributed with a mean of 0 and standard deviation of 1. The distribution of item difficulties was uniform, ranging between -2 and 2. The items in each component had difficulties distributed in the same way. The generated data sets were analysed with the RUMM2020 software (Andrich, Sheridan and Luo, 2004).

Consider the example of an assessment with five components. The assessment has 50 dichotomous items with 10 items in each component. Since there were five components of items a component analysis was also done in each condition. In the component analysis the 10 items simulated to belong to the same component were identified in the analysis as belonging to the same component. This created five items instead of 50 in the analysis. This analysis in principle takes account of the dependencies within each component.

For both response dependence and trait dependence c was varied: c=0, 0.5 and 1. In the case of response dependence with c > 0 item 2 was dependent on item 1, item 3 was dependent on item 2, item 4 was dependent on item 3, etc. In the case of trait dependence a c value of 1 results in correlations of 0.5 between the components and a c value of 0.5 results in correlations of 0.8 between the components. A c value of 0 resulted in all the components being correlated at r=1 and is the case of data fitting the dichotomous model without any violations. To study the effects of the violations of the model on the scale, analyses with the items treated as dichotomous items and as items within their components, and with the correlation values of 0, 0.8 and 1.0 were carried out.

In the Rasch models, the total score is a sufficient statistic for the person ability. Thus for each total score on a set of items, and irrespective of the pattern of correct responses across the items, there is a single ability estimate. Each total score is transformed non linearly to give the ability estimate. Thus irrespective of taking or not taking into account the dependence, the analysis will have the same raw score

range with distribution of raw scores. One concern in these studies is the range of values of the transformed scores and the degree to which violations of the model affect these values, and therefore the scale. Other related concerns investigated are the standard deviation of person locations, the person separation index, the percentage of persons below -1 standard deviation and the correlation and principal component analysis of person-item residuals.

3.1 Range and standard deviation of person locations

Figure 1 shows the range of the person locations at each level of c.

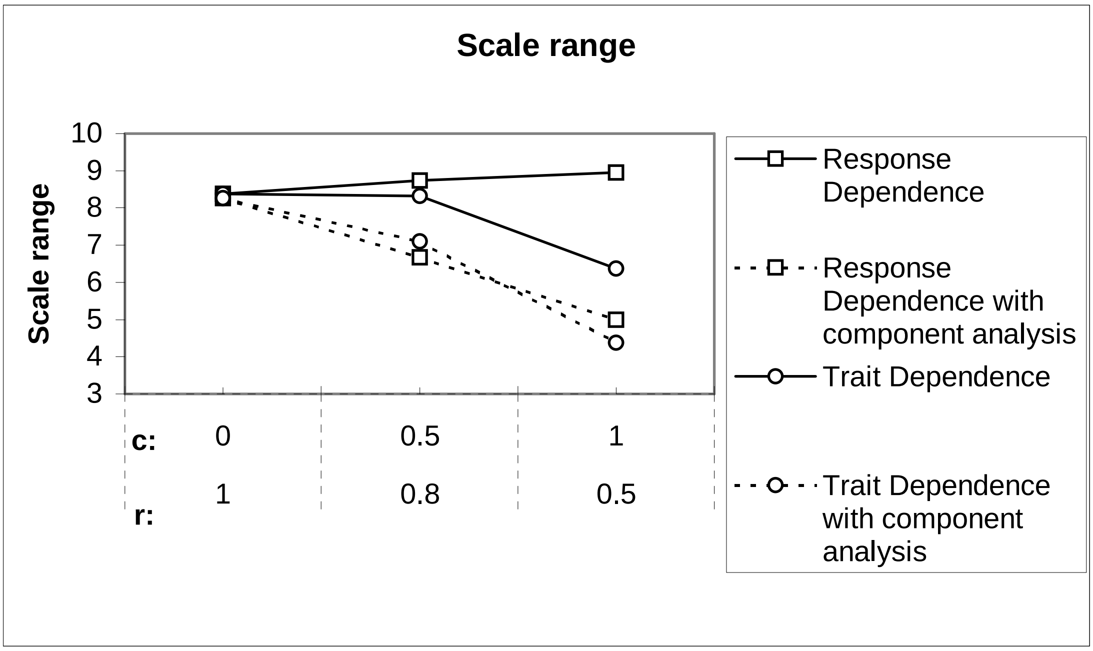

Figure 1 Scale range as a function of c.

Figure 2 shows the standard deviation of the person locations at each level of c.

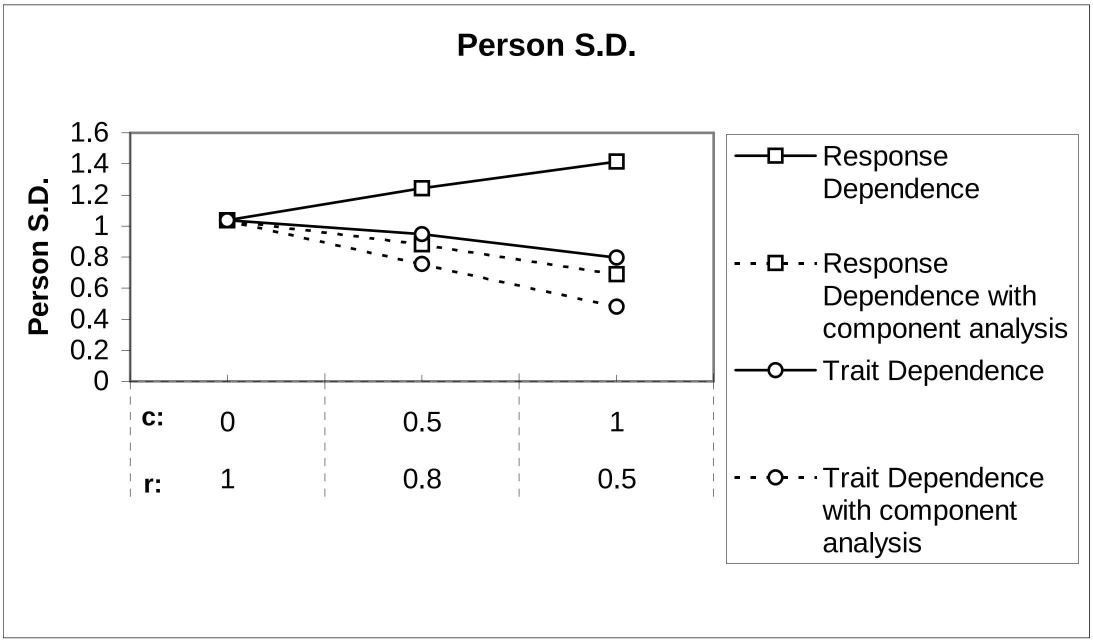

Figure 2 Standard deviation of the person location scale as a function of c.

The range is the person location values for a total score of 1 to (maximum score -1). For c=0 (no dependence) the range of the person location scale was 8.38 (i.e. -4.19 to 4.19) with a standard deviation of 1.038. Data was simulated to have a standard deviation of 1. For the response dependence condition with c=1 the scale increases to 8.96 (-4.39 to 4.57) with a standard deviation of 1.41. For the trait dependence condition (c=1, components correlated at 0.5) the scale decreases to 6.37 (-3.18 to 3.19) with a standard deviation now 0.8.

With a component analysis the scale and standard deviation remains the same when there is no dependence. However, when there is dependence (of either type) the scale and standard deviation decreases dramatically, more so with trait dependence than with response dependence.

Item standard deviation: The item standard deviation in the no dependence condition is 1.28. In the trait dependence condition (c=1) this reduces to 1.21 and in the response dependence condition (c=1) it increases to 1.57.

3.2 Person Separation

The person separation index is based on the traditional reliability formula:

$$r_{xx} = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_e^2} = \frac{\sigma_x^2 - \sigma_e^2}{\sigma_x^2},$$

where σ_x^2 is the variance of the observed total scores, σ_τ^2 is the variance of the true scores, σ_e^2 is the error variance of each measurement and $\sigma_x^2 = \sigma_\tau^2 + \sigma_e^2$. Thus the reliability is a function of both the variance of the observed estimates and the error of measurement.

A similar index is constructed with Rasch measurement, termed the person separation,

$$r_{\beta\beta} = \frac{\sigma_\beta^2}{\sigma_\beta^2 + \sigma_e^2} = \frac{\sigma_{\hat{\beta}}^2 - \sigma_{\hat{e}}^2}{\sigma_{\hat{\beta}}^2}$$

with $\hat{\sigma}_{\hat{\beta}}^2$ is the estimated variance of the locations of the persons, and $\hat{\sigma}_{\hat{e}}^2$ is the average squared standard error of measurement for each person. Again, it is a function of both the variance of the estimates scores and the error of measurement variance, and is a relevant statistic to consider in relation to specific violations of model fit.

Figure 3 shows the change in the person separation index for the four conditions.

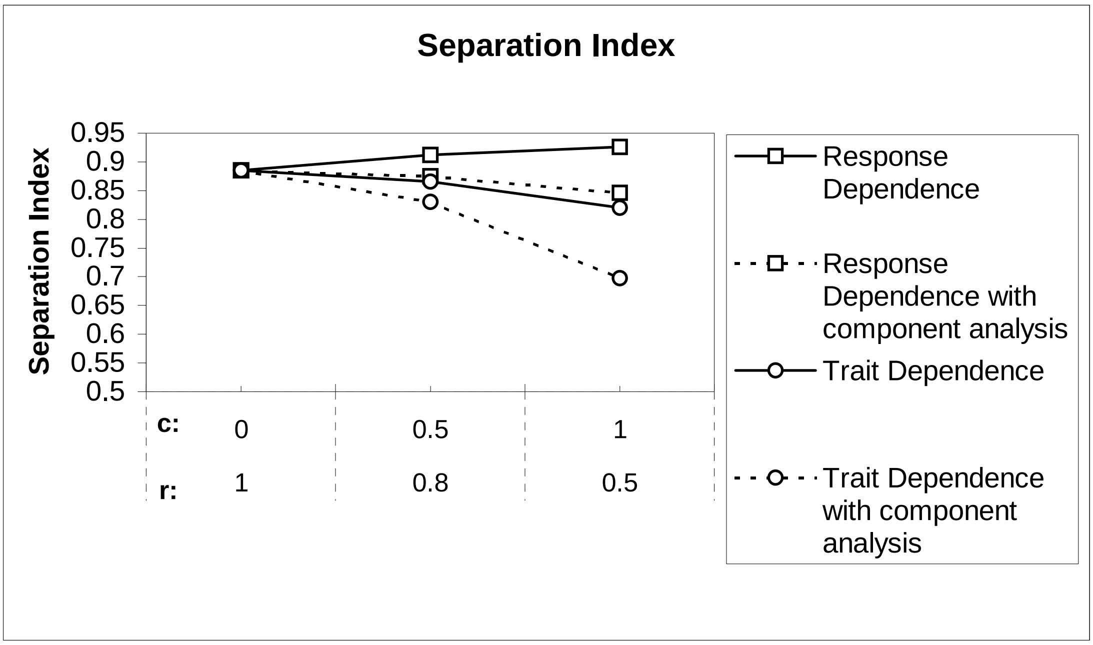

Figure 3 Person separation index as a function of c.

With no dependence the Person Separation Index is 0.89 and a component analysis does not change the index. In the response dependence condition (c=1) the index increases to 0.93 and in the trait dependence condition (c=1) the index decreases to 0.82. The index decreases in a component analysis for both types of dependence: 0.85 for response dependence (c=1) and 0.7 for trait dependence (c=1). This indicates that the index of person separation calculated when the items are placed into their components reflects the dependence in the components. This change in value can be used as an indicator of dependence. Further work is being carried out to formalise the relationship between the drop in the separation index and the degree of dependence.

3.3 Residual Principal Component analysis

A correlation and principal component analysis of person-item residuals can also locate where there might be dependencies in the responses. The results of a Principal Component analysis when there is no dependence and with c=1 for both trait and response dependence is shown in table 2. Only the results for the first 10 components are shown due to space restrictions. In the case of no dependence the percentage of variance accounted for by each component is quite similar.

Table 2 Principal Component Analysis

NO DEPENDENCE (c = 0)				
Item	**Eigen Value**	**Percent**	**Cumm Percent**	**Std Errors**
PC001	1.439	2.88%	2.88%	0.199
PC002	1.413	2.83%	5.70%	0.195
PC003	1.381	2.76%	8.47%	0.191
PC004	1.373	2.75%	11.21%	0.191
PC005	1.332	2.66%	13.88%	0.185
PC006	1.312	2.62%	16.50%	0.181
PC007	1.282	2.56%	19.06%	0.176
PC008	1.274	2.55%	21.61%	0.176
PC009	1.254	2.51%	24.12%	0.171
PC010	1.237	2.47%	26.59%	0.171
TRAIT DEPENDENCE (c = 1)				
Item	**Eigen Value**	**Percent**	**Cumm Percent**	**Std Errors**
PC001	1.896	3.79%	3.79%	0.264
PC002	1.820	3.64%	7.43%	0.254
PC003	1.734	3.47%	10.90%	0.241
PC004	1.646	3.29%	14.19%	0.227
PC005	1.356	2.71%	16.90%	0.187
PC006	1.286	2.57%	19.48%	0.178
PC007	1.276	2.55%	22.03%	0.175
PC008	1.266	2.53%	24.56%	0.175
PC009	1.241	2.48%	27.04%	0.172
PC010	1.196	2.39%	29.44%	0.164
RESPONSE DEPENDENCE (c = 1)				
Item	**Eigen Value**	**Percent**	**Cumm Percent**	**Std Errors**
PC001	2.262	4.52%	4.52%	0.313
PC002	2.205	4.41%	8.93%	0.308
PC003	2.147	4.29%	13.23%	0.300
PC004	2.066	4.13%	17.36%	0.288
PC005	1.920	3.84%	21.20%	0.267
PC006	1.784	3.57%	24.77%	0.245
PC007	1.720	3.44%	28.21%	0.239
PC008	1.669	3.34%	31.55%	0.232
PC009	1.595	3.19%	34.74%	0.220
PC010	1.511	3.02%	37.76%	0.210

3.4 Residual correlations

The residual correlation matrices for the no dependence condition as well as for the trait and response conditions with c=1 are shown in table 3. Once again, due to space restrictions correlations for the first 10 items only are shown. Correlations higher than 0.1 are highlighted. In the no dependence condition the residuals are not correlated. In the case of response dependence the dependence of item 2 on item1, 3 on 2, 4 on 3, etc. is clearly shown. In the case of trait dependence the residuals are not correlated.

Table 3 Residual matrices

NO DEPENDENCE (c = 0)

	I0001	**I0002**	**I0003**	**I0004**	**I0005**	**I0006**	**I0007**	**I0008**	**I0009**	**I0010**
I0001	1.000									
I0002	-0.035	1.000								
I0003	-0.044	-0.006	1.000							
I0004	0.012	-0.032	0.047	1.000						
I0005	0.036	0.037	-0.050	0.042	1.000					
I0006	-0.001	-0.010	-0.025	-0.037	-0.018	1.000				
I0007	-0.060	-0.081	0.039	-0.014	-0.061	-0.054	1.000			
I0008	-0.038	0.029	-0.030	-0.040	-0.020	-0.062	-0.041	1.000		
I0009	-0.068	0.015	-0.042	-0.070	-0.031	0.021	0.011	-0.009	1.000	
I0010	0.095	-0.047	-0.062	0.011	-0.013	0.021	-0.019	-0.011	-0.069	1.000

TRAIT DEPENDENCE (c = 1)

	I0001	**I0002**	**I0003**	**I0004**	**I0005**	**I0006**	**I0007**	**I0008**	**I0009**	**I0010**
I0001	1.000									
I0002	0.030	1.000								
I0003	-0.007	0.074	1.000							
I0004	0.024	0.026	0.073	1.000						
I0005	0.046	0.082	0.060	0.103	1.000					
I0006	0.063	-0.002	0.020	0.006	0.078	1.000				
I0007	0.011	-0.006	0.039	0.042	0.042	0.024	1.000			
I0008	0.061	0.047	0.027	-0.019	0.002	0.037	-0.003	1.000		
I0009	0.010	0.075	0.053	0.021	0.052	0.044	0.091	0.076	1.000	
I0010	0.057	0.002	-0.011	0.044	0.079	0.060	0.044	0.044	0.011	1.000

RESPONSE DEPENDENCE (c = 1)

	I0001	**I0002**	**I0003**	**I0004**	**I0005**	**I0006**	**I0007**	**I0008**	**I0009**	**I0010**
I0001	1.000									
I0002	0.254	1.000								
I0003	0.043	0.271	1.000							
I0004	0.004	0.082	0.372	1.000						
I0005	0.016	0.029	0.112	0.331	1.000					
I0006	0.030	0.005	0.023	0.098	0.369	1.000				
I0007	0.003	-0.065	-0.060	0.045	0.138	0.326	1.000			
I0008	-0.030	-0.017	-0.057	-0.027	0.002	0.039	0.271	1.000		
I0009	-0.008	0.008	0.005	-0.010	-0.011	-0.019	0.103	0.340	1.000	
I0010	0.076	-0.011	-0.032	-0.013	-0.035	-0.024	0.057	0.162	0.252	1.000

3.5 Effect on percentage below –1 S.D.

An important consequence of the modifications of scale is in benchmarking and identifying percentages of persons below benchmark. For the illustrative analysis, the benchmark is taken as -1.00 logits which is one standard deviation below the mean in the original persons distribution and should result in approximately 16% below the benchmark.

Figure 4 shows the percentage of persons below – 1 S.D. for different levels of c.

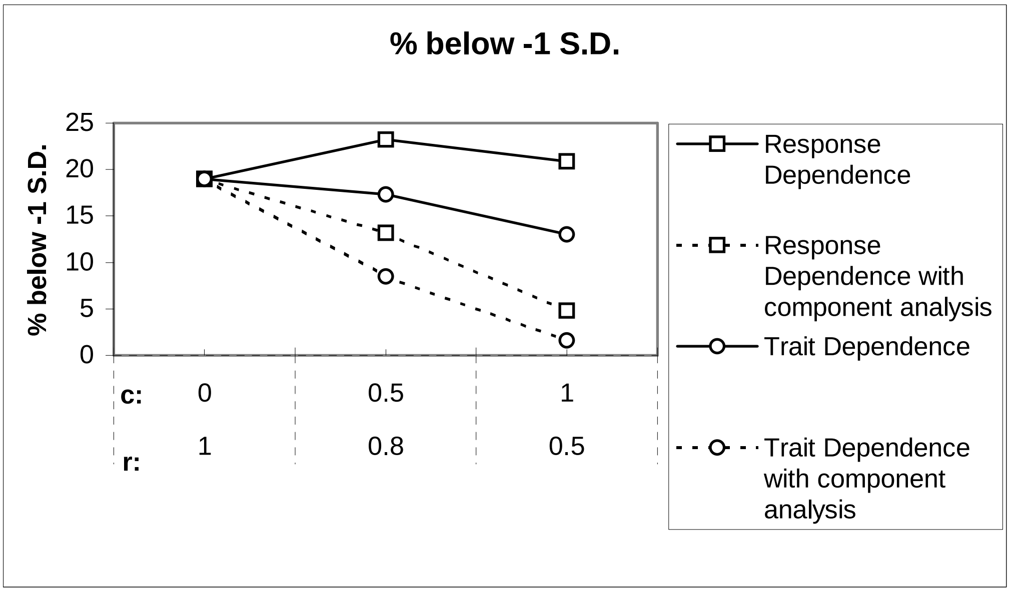

Figure 4 Percentage below –1 S.D. as a function of c.

In the case of no dependence the percentage of persons below -1 (data was simulated with a standard deviation of 1) is 19.0. For response dependence this percentage increases to 20.9 when c=1. For trait dependence the percentage decreases to 13.0 when c=1. With a component analysis the percentage decreases dramatically to 5.9 for response dependence and 3.1 for trait dependence.

4. Conclusions

Data were simulated with varying levels of trait and response dependence. When these data were analysed according to the Rasch model these violations of the model

resulted in marked effects on the unit of scale, as manifested in the range and the standard deviation of person locations. In the case when data with response dependence (as simulated in this paper) are analysed ignoring the violations of the model the scale is 'stretched' and this results in a bigger percentage of persons below -1 standard deviation. In the case when data with trait dependence are analysed ignoring the violation of the model the scale 'shrinks' resulting in a smaller percentage of persons below -1 standard deviation. In both cases when data with dependence are analysed with a component analysis the scale shrinks.

Other effects of dependence on typical statistical indices of misfit were:

The Person Separation Index increases for data with response dependence compared with data that has no dependence. The Person Separation Index decreases for data with trait dependence compared with data that has no dependence. When there is dependence the index decreases further in a component analysis (when items belonging to the same component are identified in the analysis as such).

A principal component analysis reflects some dependence, although with fairly small percentages of variance accounted for by the components.

Response dependence is clearly reflected in the residual correlation matrix. There was almost no difference between the residual correlation matrices for the trait dependence and no dependence conditions.

Although only one simulation example is reported here, the effects are robust and have been demonstrated in other examples that varied item difficulties and person abilities, numbers of items and persons, numbers of components and numbers of items in each component. Studies into dependence and its effects are continuing.

Appendix

The construction of the generating latent abilities:

To make the construction of the latent variables B_1 and B_2, which govern the responses to the items in components 1 and 2 respectively, simple and general, it is convenient to have another pair of intermediate variables A_1 and A_2. We work now with A_1 and A_2 and then define B_1 and B_2 in terms of A_1 and A_2 respectively.

Recall that β, β_1 and β_2 are standard unit normal deviates with mean 0 and variance 1.

Let $A_1 = \beta + c\beta_1$ and let $A_2 = \beta + c\beta_2$.

Because A_1 and A_2 have β in common, they will be correlated; because they have respectively β_1 and β_2 (with c>0) which are uncorrelated, they will not be correlated perfectly.

Let the correlation between the intermediate A_1 and A_2 be r_{12}. It will be shown that this is the same correlation as that between B_1 and B_2 when the latter are fully defined.

Then $r_{12} = \dfrac{\text{cov}[A_1, A_2]}{\sqrt{V[A_1]}\sqrt{V[A_2]}}$ (A1)

However, $\text{cov}[A_1, A_2] = \text{cov}[\beta + c\beta_1, \beta + c\beta_2] = \text{cov}[\beta, \beta] = V[\beta]$. This follows because the correlation among β, β_1, β_2 is mutually 0.

That is $\text{cov}[A_1, A_2] = V[\beta] = 1.$ (A2)

Now

$V[A_1] = V[\beta + c\beta_1] = V[\beta] + c^2 V[\beta_1]$ (A3)

and

$$V[A_2] = V[\beta + c\beta_2] = V[\beta] + c^2 V[\beta_2] \tag{A4}$$

and this follows again because the correlation among β, β_1, β_2 is mutually 0.

Substituting (A2), (A3) and (A4) into (A1) gives

$$r_{12} = \frac{V[\beta]}{\sqrt{V[\beta] + c^2 V[\beta_1]}\sqrt{V[\beta] + c^2 V[\beta_2]}} \tag{A5}$$

However, $V[\beta] = V[\beta_1] = V[\beta_2] = 1$.

Therefore, on simplifying (A5)

$$r_{12} = \frac{1}{1 + c^2} \tag{A6}$$

and $$c^2 = \frac{1 - r_{12}}{r_{12}} \tag{A7}$$

Clearly if $c = 0$, then $r_{12} = 1$, as it should be. The greater the value of c, the smaller the correlation.

Thus any correlation between A_1 and A_2, (and therefore between B_1 and B_2), can be defined in terms of c.

Now we define B_1 and B_2:

Define $B_1 = a_1 + \frac{b_1}{\sqrt{1 + c^2}} A_1$ and $B_2 = a_2 + \frac{b_2}{\sqrt{1 + c^2}} A_2$ (A8)

Then the means of B_1 and B_2 are respectively a_1 and a_2,

their variances are b_1^2 and b_2^2, and their intercorrelation is

$$r_{12} = \frac{1}{1+c^2}.$$

This is proved below. First note that

$$E[A_1] = E[\beta + c\beta_1] = E[\beta] + cE[\beta_1] = 0 + 0 = 0 = E[A_2]$$

and

$$V[A_1] = V[\beta + c\beta_1] = V[\beta] + c^2 V[\beta_1] = 1 + c^2 = V[A_2].$$

Then

$$E[B_1] = E[a_1 + \frac{b_1}{\sqrt{1+c^2}} A_1] = E[a_1] + E[\frac{b_1}{\sqrt{1+c^2}} A_1] = a_1 + 0 = a_1,$$

and likewise $E[B_2] = a_2$.

$$V[B_1] = V[a_1 + \frac{b_1}{\sqrt{1+c^2}} A_1] = \frac{b_1^2}{(1+c^2)} V[A_1] = \frac{b_1^2}{(1+c^2)}(1+c^2) = b_1^2$$

and likewise, $V[B_2] = b_2^2$.

Finally, $COV[B_1, B_2] = COV[a_1 + \frac{b_1}{\sqrt{1+c^2}} A_1, \quad a_2 + \frac{b_2}{\sqrt{1+c^2}} A_2]$

$$= COV[\frac{b_1}{\sqrt{1+c^2}} A_1, \quad \frac{b_2}{\sqrt{1+c^2}} A_2]$$

$$= \frac{b_1 b_2}{1+c^2} COV[A_1, A_2]$$

$$= \frac{b_1 b_2}{1+c^2}(1) = \frac{b_1 b_2}{1+c^2}$$ (from A2: $COV[A_1, A_2] = V[\beta] = 1$).

Therefore, $\dfrac{\mathrm{COV}[B_1,B_2]}{\sqrt{V[B_1]}\sqrt{V[B_2]}} = \dfrac{b_1b_2}{1+c^2}\dfrac{1}{b_1b_2} = \dfrac{1}{1+c^2} = r_{12}$ (from A6).

Special: Correlation within a component

Consider the correlation among responses within component 1:

$$r_{11} = \frac{\mathrm{cov}[A_1, A_1]}{\sqrt{V[A_1]}\sqrt{V[A_1]}} = \frac{\mathrm{cov}[A_1, A_1]}{V[A_1]} \qquad \text{(A9)}$$

$\mathrm{COV}[A_1, A_1] = \mathrm{COV}[\beta + c\,\beta_1, \beta + c\,\beta_1]$

$= E[(\beta + c\,\beta_1)(\beta + c\,\beta_1)] - E[\beta + c\,\beta_1]E[\beta + c\,\beta_1]$

$= E[\beta^2 + c\,\beta_1\,\beta + c\,\beta_1\,\beta + c^2\,\beta_1\,\beta_1] - 0 * 0$

$= E[\beta^2] + c^2\,E[\beta_1{}^2]$

$= V[\beta] + c^2\,V[\beta_1]$

and

$V[A_1] = V[\beta + c\beta_1] = V[\beta] + c^2 V[\beta_1]$

then

$$r_{11} = \frac{\mathrm{cov}[A_1, A_1]}{\sqrt{V[A_1]}\sqrt{V[A_1]}} = \frac{V[\beta] + c^2 V[\beta_1]}{V[\beta] + c^2 V[\beta_1]} = 1$$

References

Andrich, D., Sheridan, B., Luo, G. (2004). RUMM2020: A Windows interactive program for analysing data with Rasch unidimensional models for measurement. Perth Western Australia: RUMM Laboratory.