SA Journal of Industrial Psychology, 2004, 30 (4), 27-43
SA Tydskrif vir Bedryfsielkunde, 2004, 30 (4), 27-43

OVERCOMING THE EFFECTS OF DIFFERENTIAL SKEWNESS OF TEST ITEMS IN SCALE CONSTRUCTION

JOHANN M SCHEPERS
Department of Human Resource Management
Rand Afrikaans University

ABSTRACT

The principal objective of the study was to develop a procedure for overcoming the effects of differential skewness of test items in scale construction. It was shown that the degree of skewness of test items places an *upper limit* on the correlations between the items, regardless of the *contents* of the items. If the items are ordered in terms of skewness the resulting intercorrelation matrix forms a *simplex* or a *pseudo simplex*. Factoring such a matrix results in a *multiplicity* of factors, most of which are *artefacts*. A procedure for overcoming this problem was demonstrated with items from the Locus of Control Inventory (Schepers, 1995). The analyses were based on a sample of 1662 first-year university students.

OPSOMMING

Die hoofdoel van die studie was om 'n prosedure te ontwikkel om die gevolge van differensiële skeefheid van toetsitems, in skaalkonstruksie, teen te werk. Daar is getoon dat die graad van skeefheid van toetsitems 'n *boonste grens* plaas op die korrelasies tussen die items ongeag die *inhoud* daarvan. Indien die items gerangskik word volgens graad van skeefheid, sal die interkorrelasiematriks van die items 'n *simpleks* of *pseudosimpleks* vorm. Indien so 'n matriks aan faktorontleding onderwerp word, lei dit tot 'n *veelheid* van faktore waarvan die meerderheid *artefakte* is. 'n Prosedure om hierdie probleem te bowe te kom, is gedemonstreer met behulp van die items van die Lokus van Beheer-vraelys (Schepers, 1995). Die ontledings is op 'n steekproef van 1662 eerstejaaruniversiteitstudente gebaseer.

In the construction of psychometric tests it is normally assumed that the items included in a test form a *linear scale*. To test this assumption the *dimensionality* of the vector space of test items must first be determined. Should it turn out to be *multidimensional*, the test items must first be categorised according to the *construct* measured. The categorisation of the test items can be done with the aid of factor analysis, but the procedure is not free of problems.

The gist of the problem concerns the fact that test items usually vary in terms of their *degree of* skewness (i.e. are differentially skew), and this affects their mutual intercorrelations. If test items that are differentially skew are subjected to factor analysis, a *multiplicity* of factors is usually obtained. This tends to obscure the true structure of the intercorrelation matrix of test items (cf. Ferguson, 1941).

To illustrate the problem, *binary* items (i.e. items scored dichotomously), with different *marginal splits*, will be used. Thereafter the principle will be generalised to items where the responses are endorsed on *continuous* scales (usually five or seven-point scales) which are often differentially skew. But to start off with the concept 'marginal split' will first be clarified:

If test items are scored dichotomously, the *proportion* of respondents endorsing an item according to the *scoring key* (p_g) and the proportion of respondents not endorsing the item according to the scoring key (q_g), will vary from item to item. The ratio of p_g to q_g is usually referred to as the *marginal split* of item $_g$. In the general literature p_g is usually referred to as the '*difficulty value*' of item $_g$. For convenience sake, the expression 'p-value' will be used in this paper.

In Table 1 the maximum intercorrelations of 10 test items with different *marginal splits* are given. The items have been arranged according to their marginal distributions from 0,05:0,95 to 0,95:0,05.

From Table 1 it is evident that there is a clear *gradient* underlying the intercorrelations: adjacent items correlate the highest with one another, for example items 1 and 2, whilst items that are far apart, for example items 1 and 10, correlate the lowest. The highest intercorrelations are adjacent to the principal diagonal and systematically taper off as you move away from the principal diagonal. This is true for both the rows and the columns of the intercorrelation matrix. Such intercorrelation matrices are known as *simplexes*, and must not be confused with ordinary *unidimensional* intercorrelation matrices.

The *inverse* of a simplex has rather special properties: the *principal diagonal* is *positive* and the two *adjacent diagonals* are *negative*. All the other elements are equal to zero (Guttman, 1954, 1955, 1957; Schepers, 1962; Jöreskog, 1970).

The inverse of the intercorrelation matrix in Table 1 is given in Table 2.

From Table 2 it is clear that the inverse of the intercorrelation matrix, in Table 1, reveals all the properties of a simplex. The implications of this will now be carefully scrutinised.

The significance of a simplex structure is clearest if the standardised *multiple regression coefficients* for predicting each variable (item) by the rest, are viewed. In the case of a *perfect simplex* only adjacent variables can be used to predict a particular variable. The regression coefficients of all the other variables are equal to zero.

The regression coefficients for predicting each variable by the rest, in respect of Table 1, are given in Table 3 (cf. Schepers, 1962, p. 301).

From Table 3 it can be seen that the regression coefficients in the two diagonals adjacent to the principal diagonal vary from 0,3052 to 0,6063. All the other coefficients are equal to zero.

As mentioned earlier, the intercorrelation matrix in Table 1 represents the *maximum intercorrelations* of 10 test items with specified *marginal distributions*. Furthermore, the assumption was made that all the items relate to a single central construct (Bohrnstedt & Knoke, 1988; Gorsuch, 1974; Guilford, 1950; Magnusson, 1967). It would therefore be logical to expect a single factor underlying the matrix of intercorrelations. The structure of the intercorrelation matrix was accordingly investigated.

The eigenvalues of the unreduced intercorrelation matrix are given in Table 4.

Requests for copies should be addressed to: JM Schepers, Department of Human Resource Management, RAU, PO Box 524, Auckland Park, 2006

Table 1
Intercorrelation matrix of test items with different degrees of skewness

	1	2	3	4	5	6	7	8	9	10
1	1,0000	0,5461	0,3974	0,3126	0,2536	0,2075	0,1683	0,1325	0,0964	0,0526
2	0,5461	1,0000	0,7276	0,5725	0,4644	0,3800	0,3083	0,2425	0,1765	0,0964
3	0,3974	0,7276	1,0000	0,7868	0,6383	0,5222	0,4237	0,3333	0,2425	0,1325
4	0,3126	0,5725	0,7868	1,0000	0,8112	0,6637	0,5385	0,4237	0,3083	0,1683
5	0,2536	0,4644	0,6383	0,8112	1,0000	0,8182	0,6637	0,5222	0,3800	0,2075
6	0,2075	0,3800	0,5222	0,6637	0,8182	1,0000	0,8112	0,6383	0,4644	0,2536
7	0,1683	0,3083	0,4237	0,5385	0,6637	0,8112	1,0000	0,7868	0,5725	0,3126
8	0,1325	0,2425	0,3333	0,4237	0,5222	0,6383	0,7868	1,0000	0,7276	0,3974
9	0,0964	0,1765	0,2425	0,3083	0,3800	0,4644	0,5725	0,7276	1,0000	0,5461
10	0,0526	0,0964	0,1325	0,1683	0,2075	0,2536	0,3126	0,3974	0,5461	1,0000
s_g	0,2179	0,3571	0,4330	0,4770	0,4975	0,4975	0,4770	0,4330	0,3571	0,2179
p_g	0,05	0,15	0,25	0,35	0,45	0,55	0,65	0,75	0,85	0,95
q_g	0,95	0,85	0,75	0,65	0,55	0,45	0,35	0,25	0,15	0,05

Note: $s_g = \sqrt{p_g q_g}$, $r_{phi} = \dfrac{p_{gk} - p_g p_k}{s_g s_k}$

Table 2
Inverse of intercorrelation matrix

	1	2	3	4	5	6	7	8	9	10
1	**1,4250**	-0,7782								
2	-0,7782	**2,5500**	-1,5462							
3		-1,5462	**3,7500**	-2,0653						
4			-2,0653	**4,5500**	-2,3729					
5				-2,3729	**4,9500**	-2,4750				
6					-2,4750	**4,9500**	-2,3729			
7						-2,3729	**4,5500**	-2,0653		
8							-2,0653	**3,7500**	-1,5462	
9								-1,5462	**2,5500**	-0,7782
10									-0,7782	**1,4250**

Table 3
Standardised multiple regression coefficients for predicting each item by the rest

	1	2	3	4	5	6	7	8	9	10
1	-	**0,5461**								
2	**0,3052**	-	**0,6063**							
3		**0,4123**	-	**0,5508**						
4			**0,4539**	-	**0,5215**					
5				**0,4794**	-	**0,5000**				
6					**0,5000**	-	**0,4794**			
7						**0,5215**	-	**0,4539**		
8							**0,5508**	-	**0,4123**	
9								**0,6063**	-	**0,3052**
10									**0,5461**	

TABLE 4
EIGENVALUES OF INTERCORRELATION MATRIX

Root	Eigenvalue
1	**5,0752**
2	**1,7929**
3	**1,0005**
4	0,6704
5	0,4671
6	0,3281
7	0,2374
8	0,1785
9	0,1389
10	0,1111
Trace	10,000

From Table 4 it is clear that there are *three* eigenvalues greater than unity. Accordingly three factors were postulated (Kaiser, 1961).

Next, the intercorrelation matrix was subjected to a principal factor analysis. Three factors were extracted and rotated to simple structure by means of a Varimax rotation. The rotated factor matrix is given in Table 5.

TABLE 5
ROTATED FACTOR MATRIX (VARIMAX)

Variable	Factor 1	Factor 2	Factor 3	h_j^2
1	0,071	**0,545**	0,060	0,3056
2	0,192	**0,874**	0,092	0,8088
3	0,464	**0,730**	0,083	0,7553
4	**0,676**	0,542	0,114	0,7643
5	**0,822**	0,347	0,192	0,8322
6	**0,822**	0,192	0,347	0,8322
7	**0,676**	0,114	0,542	0,7643
8	0,464	0,083	**0,730**	0,7553
9	0,192	0,092	**0,874**	0,8088
10	0,071	0,060	**0,545**	0,3056

Table 5 shows that items 4, 5, 6 and 7 have high loadings on Factor 1, whilst items 1, 2 and 3 have high loadings on Factor 2. Items 8, 9 and 10 have high loadings on Factor 3. Items 1, 2 and 3 have p-values between 0,05 and 0,25; items 4, 5, 6 and 7 have p-values between 0,35 and 0,65, and items 8, 9 and 10 have p-values between 0,75 and 0,95. The items therefore *cluster* according to their *marginal distributions*.

It is also striking that the *communalities* of items 1 and 10 are considerably lower than the rest. This phenomenon is typical of simplexes, because in a simplex it is only adjacent items that share *common variance*, and the *first* and the *last* items have only one adjacent item each.

All the limitations that have been referred to above are the direct consequence of the simplex structure of the particular intercorrelation matrix. But the most troublesome aspect is the *artefactual factors* that emerge if a simplex is subjected to a principal factor analysis. In the present case there were only three factors, but it is not uncommon to obtain a *multiplicity of factors*, particularly in respect of large matrices.

Different solutions have already been proposed to overcome the problem of *artefactual factors*. Gorsuch (1974, p. 262), for instance, maintains that the best solution is to avoid working with variables that are *too skew*. His suggestion is highly acceptable, provided the variables used derive from well constructed measuring instruments. Where necessary the obtained scores can even be *normalised* before doing a factor analysis. But when working with test items his suggestion is not acceptable. The *discrimination power* and *reliability* of a test will be lowered if all the items have equal p-values. A wide distribution of p-values (0,10 to 0,90) is necessary to ensure good discrimination power and reliability. The rationale for this should be apparent from an analysis of Kuder-Richardson Formula 20 (Kuder & Richardson, 1937). The derivation of the formula is given in Appendix 1.

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{\mu - \dfrac{\mu^2}{K} - K\sigma_p^2}{\sigma_x^2}\right], \text{ where} \qquad (1)$$

K = number of test items

m = mean of test

p_g = proportion of subjects endorsing item g according to the key

σ_p^2 = variance of p_g's

The larger $K\sigma_p^2$, the larger the test reliability would be (Horst, 1953).

Tucker (1949, p. 119) has shown that the *variance* of the p-values for a *rectangular distribution* of the p_g's, varying from 0,00 to 1,00, is equal to 0,083. For a *normal distribution* with p = 0,00 at -3σ and p = 1,00 at +3σ, the variance is equal to 0,028.

From the foregoing it is clear that for a longer test or scale, with a wide distribution of p_g's, the value of $K\sigma_p^2$ can be quite substantial.

Formula 1 applies only to *binary data*.

For *continuous data*, Kuder-Richardson Formula 20 and Cronbach's coefficient alpha (Cronbach, 1951), can be used. The transformed KR_{20} formula can be written as follows:

$$KR_{20} = \left[1 - \frac{K\sum_{g=1}^{k}\sigma_g^2 - \sigma_x^2}{(K-1)\sigma_x^2}\right], \qquad (2)$$

where

K = number of test items

$\sum_{g=1}^{K}\sigma_g^2$ = sum of item variances

σ_x^2 = test variance

[The derivation of Formula 2 is given in Appendix 1].

The total test variance can never exceed $K\sum_{g=1}^{K}\sigma_g^2$ (cf. Schepers, 1992, p. 33), therefore, the greater the test variance, the *higher* the *reliability* will be.

From the foregoing it is clear that Gorsuch's suggestion of avoiding items that are too skew, will lead to a reduction in test reliability.

Horst (1965, p. 516) maintained that one way of getting rid of the artefactual factors is to fit an appropriate simplex to the binary data matrix and to separate it from the true structure. This is equivalent to throwing the baby away with the bath water.

The fitting of a simplex poses problems of its own. Firstly, an appropriate model must be decided upon, and then an appropriate method must be found to fit the model to the data. If the measuring units of all the variables are the *same*, a simplex can be fitted to the *variance-covariance matrix*, but if the units of measurement *differ*, the *intercorrelation matrix* must be used (Jöreskog, 1970).

Jöreskog (1970, p. 122) distinguishes six different simplex models. These models can broadly be divided into Markov simplexes and Wiener simplexes. The Markov simplexes can be subdivided into perfect simplexes, quasi-simplexes and restricted quasi-simplexes. Similarly, Wiener simplexes can be subdivided into perfect simplexes, quasi-simplexes and quasi-simplexes with *equal error variances*. Markov simplexes are *scale free*, i.e. the measuring units of the variables do not have to be the same. Wiener simplexes, on the other hand, are *scale dependent*, ie. they can only be used if the units of measurement of all the variables are the same (Jöreskog, 1970, p. 128).

Perfect simplexes only arise if *true scores* are used. The moment there is *measurement error* it becomes a quasi-simplex. According to Jöreskog (1970, p. 130) the fitting of a perfect simplex is straightforward, but the fitting of quasi-simplexes is complex and requires iterative procedures. Reasonable fits can be obtained with the LISREL-program (Jörgeskog & Sorbom, 1982).

There are many intercorrelation matrices that superficially resemble simplexes, but that are really *pseudo-simplexes*. This complicates the issue of fitting simplexes even further. It is also doubtful whether the fitting of a simplex overcomes the problem of differential skewness, and the generation of 'difficulty factors'.

So far the research problem has been developed in terms of *binary* items that are differentially skew. It will now be broadened to *continuous* items that are differentially skew.

The principal objectives of the present study were to:
1. Examine the *gradients* of the correlations (by row and by column) in an intercorrelation matrix based on *continuous* variables, arranged in terms of degree of *skewness*.
2. Determine the *factor structure* of an intercorrelation matrix based on *continuous variables*, arranged in terms of degree of *skewness*.
3. *Develop* and *evaluate* a procedure for overcoming the effects of differential skewness on the factor structure of test items, and to examine the properties of the resulting scales.

METHOD

Sample
The full complement of first-year university students at the Rand Afrikaans University, during 1995, was subjected to an extensive psychometric test programme. The programme stretched over four days, and to ensure complete records, students who did not attend all the test sessions were excluded from the sample. The final sample consisted of 1662 students, and can be considered *representative* of the population of first-year university students at the Rand Afrikaans University, during 1995. The ages of the students varied from 26 to 54 years, with a mean of 27,30 years and a standard deviation of 1,842 years. As far as gender is concerned 49,8% were female and 47,2% were male. Missing information accounted for 3,0%. The majority of the students were Afrikaans-speaking (969). Three hundred and seventy nine were English-speaking, and 195 spoke both English and Afrikaans. Only 27 had an African language as vernacular. Thirty-nine spoke other languages, and 53 did not indicate their home language. As far as ethnic group is concerned 88,7% were White, 1,4% were Indian, 4,7% were Coloured and 2,2% were African.

Measuring instrument
For the purpose of this study the items of the Locus of Control Inventory (Schepers, 1995), were used.

The construct of 'locus of control' was created by Rotter (1966) and pertains to a person's expectation of reinforcement of his/her behaviour, arising from the social environment. Therefore, it is theoretically based on social learning theory (Mischel, 1979). Rotter (1966) distinguished between two different orientations in people, namely an *internal control orientation* and an external control orientation. People with an internal control orientation are convinced that their behaviour depends on their own achievements, abilities and dedication, whereas people with an external control orientation believe that random or fortuitous events, fate, Lady Luck and certain influential people are responsible for their behaviour.

Conceptually, the Locus of Control Inventory (LCI) is based on attribution theory and social learning theory. People are constantly attuned to finding the causes of their behaviour and those of others. The *attributing* of causes to specific behaviour is called *attributions*. The causative attributions that people make and their interpretation thereof determine their perceptions of the social world to a large extent. Is it a friendly or a hostile world? Is it a just or unjust world? Is it a predictable or an unpredictable world? Can we exercise control over certain events through our own abilities or are our lives controlled by certain influential people? According to attribution theory the causes of human behaviour can be divided into two broad categories, namely *dispositional* causes and *situational* causes. Dispositional causes pertain to one's natural disposition and include the organismic attributes of people. Situational causes pertain to the external world and include all environmental factors (Roediger III et al., 1991).

Social learning theory links up with attribution theory: whereas social learning theory deals with the nature of reinforcements arising from the social behaviour of the learner, attribution theory pertains to the way in which a person gathers information about the stable or invariant characteristics of others – their motives, intentions and traits – as well as the external world (Baron, Byrne & Kantowitz, 1980).

A construct closely related to internal control is *autonomy*. Autonomy can be defined as 'the tendency to attempt to master or be effective in the environment, to impose one's wishes and designs on it' (Wolman, 1973, p. 37). It is expected that persons high on autonomy would seek control of situations that offer possibilities of change, would readily accept the challenge of solving complex problems, would take the initiative in situations requiring leadership, would prefer to work on their own and to structure their own work programme.

With attribution theory and social learning theory as frames of reference, the domain of locus of control was extensively sampled. Altogether 80 items were written, representing the constructs of internal control, external control and autonomy. Roughly equal numbers of items were written in respect of each of the constructs.

The items of the LCI are all in the form of *questions* and the responses are endorsed on a seven-point scale. Only the end-points of the scales are verbally anchored. Separate answer sheets that can be read by an optical page reader are used, and the responses can be read directly onto a stiffy or compact disk.

Procedure
The LCI was applied to the full complement of first-year university students at the Rand Afrikaans University during 1995. Thereafter the answer sheets were carefully scrutinised for double markings or incompleteness. Where necessary the markings were made clearer and stray marks were erased. If more than two items were left blank or spoilt, the person's record was not used. If one or two items were left blank or spoilt, the person's mean for that *construct* (according to the a priori key) was estimated, and his/her item mean was substituted for the particular item. Care was taken not to estimate *item means* if several respondents skipped the same item. In this way 1662 complete records were obtained.

Statistical analysis
In order to examine the *gradients* of the correlations in an intercorrelation matrix of *continuous* variables that have been arranged according to their degree of *skewness*, the means, standard deviations, coefficients of skewness and kurtosis of the 80 items of the LCI were computed. Next, the items were arranged according to their degree of skewness. To ensure stability of the data, composite scores were formed by adding the scores of *eight adjacent* items together. In this way *10* new variables were formed The *10* new variables were then intercorrelated. Following this the *inverse* of the obtained intercorrelation matrix was computed.

To determine the *factor structure* of an intercorrelation matrix based on *continuous* variables arranged in terms of degree of skewness, the same data as above were used, except that the subscores were based on *parcels* of *four adjacent* items instead of eight. The resulting 20 subscores were then intercorrelated and subjected to a principal factor analysis. The obtained factor matrix was rotated to simple structure by means of a Direct Oblimin rotation.

To overcome the effects of differential skewness of the items of the LCI, the following procedure (as suggested in Schepers, 1992) was followed:
1. The 80 items were intercorrelated.
2. The eigenvalues of the unreduced intercorrelation matrix were calculated.
3. As many factors as there were eigenvalues greater than unity were postulated.
4. An iterative principal factor analysis was done.
5. Iteration was done on the number of factors as determined at step 3.
6. The obtained factor matrix was rotated to simple structure by means of a Varimax rotation.
7. All the items with high negative loadings were reflected.
8. All the items with high loadings on a specific factor were added together and a subscore for each factor was computed. Every item was used only once.
9. The obtained subscores were intercorrelated and steps 2 to 4 were repeated.
10. The obtained factor matrix was rotated to *simple structure* by means of a Direct Oblimin rotation.
11. All subscores with negative loadings on the *first principal axis* were reflected.
12. Separate *scales* were formed, corresponding to each of the factors, by grouping all the items together that had substantial loadings on a factor, ie. all the items in the relevant subscores (cf. step 8).
13. Separate item analyses (NP50) were done for each of the scales formed.
14. Iteration was done in terms of the *indices of reliability* of the test items.
15. The *reliability* of the scales were determined by means of Cronbach's coefficient alpha.

RESULTS

Objective 1: Gradients of correlations in an intercorrelation matrix based on continuous variables, arranged in terms of skewness

As a first step, the means, standard deviations, and coefficients of skewness of the 80 items of the LCI were computed. Next, the items were arranged according to their degree of skewness. The descriptive statistics are given in Table 6.

TABLE 6
MEANS, STANDARD DEVIATIONS AND COEFFICIENTS OF SKEWNESS OF THE TEST ITEMS

	Variable	Mean	Standard deviation	Coefficient of skewness
New variable 1	Q19	6,3285	0,8537	-1,7592
	Q60	6,1330	1,1200	-1,6226
	Q61	5,9675	1,1309	-1,6141
	Q42	6,1005	1,0588	-1,5781
	Q37	5,9537	1,0426	-1,4967
	Q18	6,0632	1,0177	-1,4407
	Q49	6,2202	0,8993	-14065
	Q10	6,1258	0,9300	-1,2905
New variable 2	Q63	6,1943	0,9003	-1,2590
	Q75	5,9693	1,0867	-1,2468
	Q59	5,7443	1,3507	1,2163
	Q31	5,8207	1,1219	-1,1491
	Q13	5,7966	0,9513	-1,0591
	Q62	5,6227	1,2052	-1,0363
	Q28	5,4633	1,3655	-1,0208
	Q8	5,6071	1,1489	-1,0048
New variable 3	Q22	5,8039	1,1023	-0,9947
	Q16	5,1799	1,6189	-0,9934
	Q67	5,7635	1,0554	-0,9602
	Q66	6,0897	0,9021	-0,9457
	Q69	5,7419	1,1072	-0,9350
	Q48	5,3538	1,3030	-0,9026
	Q24	5,4398	1,2491	-0,8827
	Q54	5,4266	1,1774	-0,8541
New variable 4	Q40	5,3057	1,3288	-0,8485
	Q2	5,4940	1,2772	-0,8381
	Q73	5,1986	1,2821	-0,8258
	Q70	5,6492	1,2166	-0,8215
	Q6	5,6859	1,0957	-0,8036
	Q68	5,7389	1,0071	-0,7734
	Q46	5,3995	1,1133	-0,7371
	Q76	5,0235	1,5189	-0,7322
New variable 5	Q55	5,5890	1,0004	-0,7082
	Q27	5,6227	1,0881	-0,6982
	Q7	5,6637	1,0453	-0,6947
	Q17	5,3153	1,2883	-0,6697
	Q29	5,2882	1,2113	-0,6390
	Q32	5,2515	1,2339	-0,6214
	Q30	5,2058	1,3249	-0,6133
	Q25	5,2744	1,3027	-0,6092
New variable 6	Q14	5,2010	1,1651	-0,6063
	Q74	5,2100	1,1291	-0,6023
	Q5	5,2557	1,0171	-0,5963
	Q44	5,1871	1,0226	-0,5259
	Q39	4,7786	1,3998	-0,5103
	Q9	5,0391	1,2906	-0,5086
	Q26	4,6546	1,5714	-0,4864
	Q1	4,7960	1,3723	-0,4761

New variable 7	Q64	4,3454	1,2960	-0,4112
	Q50	4,7569	1,3762	-0,3913
	Q47	4,6191	1,4449	-0,3822
	Q15	4,6907	1,5028	-0,3266
	Q3	4,5000	1,3634	-0,3011
	Q77	4,3063	1,7838	-0,2711
	Q71	4,4302	1,2917	-0,1796
	Q51	3,6516	1,3752	-0,0815
New variable 8	Q72	4,2942	1,3370	-0,0414
	Q65	3,7100	1,6761	-0,0330
	Q80	3,6342	1,5409	0,0431
	Q38	3,5812	1,4367	0,1515
	Q20	33965	1,3670	0,1660
	Q4	3,6203	1,8354	0,1662
	Q21	3,4260	1,3251	0,1872
	Q57	3,2816	1,4653	0,2859
New variable 9	Q36	3,1949	1,4547	0,2941
	Q43	3,1270	1,6261	0,4077
	Q56	3,0355	1,4983	0,5286
	Q35	2,8670	1,5248	0,5352
	Q33	2,8008	1,4597	0,5819
	Q12	2,7882	1,5176	0,5919
	Q23	2,9777	1,4437	0,6314
	Q79	2,6901	1,4626	0,6807
New variable 10	Q34	2,6444	1,5183	0,7148
	Q78	2,9362	1,6902	0,7495
	Q41	2,5487	1,4407	0,8386
	Q45	2,5048	1,3590	0,9770
	Q11	2,4711	1,3919	0,9944
	Q58	2,3424	1,5642	1,1727
	Q53	2,1721	1,3036	1,2429
	Q52	2,1420	1,4670	1,4317

Note: Items arranged according to degree of skewness.

From Table 6 it can be seen that the coefficients of skewness vary from –1,7592 to 1,4317.

Next, new variables were formed by adding the scores of eight *adjacent* items together. The new variables were then arranged in terms of their degree of skewness. The descriptive statistics of the new variables are given in Table 7.

TABLE 7
MMEANS, STANDARD DEVIATIONS AND COEFFICIENTS OF SKEWNESS AND KURTOSIS OF THE NEW VARIABLES

	Variable	Standard deviation	Coefficient of skewness	Coefficient of kurtosis
New variable 1	48,892	4,750	-1,006	1,605
New variable 2	46,218	4,677	-0,421	0,109
New variable 3	44,799	5,102	-0,352	0,397
New variable 5	43,211	5,029	-0,343	0,046
New variable 4	43,495	4,962	-0,233	0,072
New variable 6	40,122	4,724	-0,065	-0,019
New variable 7	35,300	4,534	-0,054	0,189
New variable 8	28,944	5,453	0,012	-0,176
New variable 9	23,481	6,441	0,276	-0,254
New variable 10	19,762	6,219	0,478	0,125

Note: Standard error of coeffient of skewness = 0,06

Standard error of coefficient of kurtosis = 0,12

From Table 7 it can be seen that the coefficients of skewness range from –1,006 to 0,478. The new variables are therefore less skew than the original items.

Next, the new variables were intercorrelated. The matrix of intercorrelations is given in Table 8.

Table 8 shows that the highest correlations are *adjacent* to the *principal diagonal* and taper off as you move from left to right and from top to bottom. However, small departures from this trend is also visible. These gradients are typical of a simplex. Furthermore, the column totals also reveal the typical pattern of a simplex: the successive totals *increase* in size until a *maximum* is reached, and then systematically decrease in size. However, a more objective test is to inspect the *inverse* of the correlation matrix.

The *inverse* of the intercorrelation matrix was accordingly calculated and is given in Table 9.

It is evident from Table 9 that the principal diagonal is *positive* and the two adjacent diagonals are *negative*. However, the *off-diagonals* are not equal to *zero*. The intercorrelation matrix is therefore a *pseudo-simplex*.

From the foregoing it is clear that if *continuous* variables are arranged according to their degree of *skewness*, the correlations between the variables show *gradients* similar to that of a simplex. It is therefore expected that such matrices will generate *factors of skewness*, if factor analysed.

Objective 2: Factor structure of an intercorrelation matrix based on continuous variables, arranged in terms of degree of skewness

To determine the *factor structure* of an intercorrelation matrix based on *continuous* variables, arranged according to degree of *skewness*, the LCI data were used. Parcels of *four* adjacent items in terms of skewness, were formed. The descriptive statistics are given in Table 10.

From Table 10 it is clear that the coefficients of skewness vary from –1,16 to 0,77.

Next, the new variables were intercorrelated. The matrix of intercorrelations is given in Table 11.

Although there is not a clear gradient visible in Table 11, the correlations nevertheless become smaller as you move from left to right and from top to bottom.

The eigenvalues of the unreduced intercorrelation matrix are given in Table 12.

From Table 12 it is evident that there are *three* eigenvalues greater than unity, suggesting three factors.

Three factors were extracted and rotated to simple structure by means of a Direct Oblimin rotation. The rotated factor matrix is given in Table 13.

From Table 13 it is clear that parcels 1, 2, 3 and 9 have high loadings on Factor III. Parcels 4 to 14 (excluding parcel 9) load on Factor I, and parcels 15 to 20 load on Factor II. Each of the factors have high loadings on parcels with fairly similar coefficients of skewness. The only exception being parcel 9. It is therefore reasonable to identify the three factors as *factors of skewness*.

In order to examine the *contents* of the three factors obtained, the items represented by the *subscores* were listed. The listing of the items is given in Table 14.

Table 8
Matrix of intercorrelations of the new variables arranged according to their degree of skewness

	New var 1	New var 2	New var 3	New var 5	New var 4	New var 6	New var 7	New var 8	New var 9	New var 10	Total
New var 1	1,000	0,640	0,539	0,548	0,535	0,386	0,174	-0,152	-0,157	-0,303	3,211
New var 2	0,640	1,000	0,606	0,602	0,600	0,533	0,224	-0,214	-0,237	-0,379	3,375
New var 3	0,539	0,606	1,000	0,670	0,660	0,554	0,280	-0,158	-0,150	-0,270	3,732
New var 5	0,548	0,602	0,670	1,000	0,615	0,565	0,266	-0,134	-0,078	-0,253	3,801
New var 4	0,535	0,600	0,660	0,615	1,000	0,556	0,258	-0,210	-0,193	-0,332	3,488
New var 6	0,386	0,533	0,554	0,565	0,556	1,000	0,351	-0,121	-0,130	-0,233	3,461
New var 7	0,174	0,224	0,280	0,266	0,258	0,351	1,000	0,230	0,158	0,151	3,093
New var 8	-0,152	-0,214	-0,158	-0,134	-0,210	-0,121	0,230	1,000	0,444	0,515	1,201
New var 9	-0,157	-0,237	-0,150	-0,078	-0,193	-0,130	0,158	0,444	1,000	0,663	1,319
New var 10	-0,303	-0,379	-0,270	-0,253	-0,332	-0,233	0,151	0,515	0,663	1,000	0,559
Total	3,211	3,375	3,732	3,801	3,488	3,461	3,093	1,201	1,319	0,559	

Table 9
Inverse of intercorrelation matrix

	New var 1	New var 2	New var 3	New var 5	New var 4	New var 6	New var 7	New var 8	New var 9	New var 10
New var 1	1,909	-0,783	-0,230	-0,322	-0,260	0,156	0,001	-0,045	-0,050	0,144
New var 2	-0,783	2,390	-0,316	-0,331	-0,271	-0,348	-0,095	0,046	0,099	0,253
New var 3	-0,230	-0,316	2,414	-0,719	-0,652	-0,263	-0,127	0,036	0,064	-0,039
New var 5	-0,322	-0,331	-0,719	2,333	-0,357	-0,411	-0,034	0,010	-0,265	0,135
New var 4	-0,260	-0,271	-0,652	-0,357	2,266	-0,349	-0,133	0,120	0,016	0,171
New var 6	0,156	-0,348	-0,263	-0,411	-0,349	1,842	-0,346	0,030	0,044	0,060
New var 7	0,001	-0,095	-0,127	-0,034	-0,133	-0,346	1,328	-0,281	-0,051	-0,225
New var 8	-0,045	0,046	0,036	0,010	0,120	0,030	-0,281	1,463	-0,247	-0,484
New var 9	-0,050	0,099	0,064	-0,265	0,016	0,044	-0,051	-0,247	1,871	-1,117
New var 10	0,144	0,253	-0,039	0,135	0,171	0,060	-0,225	-0,484	-1,117	2,257

Table 10
Means, standard deviations and coefficients of skewness and kurtosis of the new variables

	Variable	Standard deviation	Coefficient of skewness	Coefficient of kurtosis
New variable 2	24,36	2,61	-1,16	2,54
New variable 1	24,53	2,74	-0,92	0,83
New variable 3	23,73	2,82	-0,70	0,86
New variable 5	22,84	3,12	-0,54	0,29
New variable 10	21,02	3,14	-0,45	0,22
New variable 4	22,49	2,79	-0,43	-0,03
New variable 7	21,65	3,05	-0,38	0,15
New variable 6	21,96	2,89	-0,34	0,23
New variable 9	22,19	2,78	-0,33	-0,09
New variable 11	20,85	3,03	-0,33	0,10
New variable 8	21,85	2,81	-0,20	-0,14
New variable 14	16,89	2,87	-0,12	0,13
New variable 13	18,41	2,98	-0,10	0,01
New variable 12	19,27	2,81	0,01	-0,05
New variable 15	15,22	3,04	0,01	0,16
New variable 16	13,72	3,72	0,04	-0,48
New variable 17	12,22	3,94	0,21	-0,48
New variable 19	10,63	3,84	0,36	-0,19
New variable 18	11,26	3,57	0,37	0,03
New variable 20	9,13	3,65	0,77	0,38

Note: Standard error of coefficient of skewness = 0,06
Standard error of coefficient of kurtosis = 0,12

From Table 14 it is clear that 24 of the 40 items loading on Factor I, are classified as *Autonomy* according to the scoring key. Eleven of the items are classified as *Internal Control* and five are classified as *External Control*.

As far as Factor II is concerned 23 of the 24 items are classified as *External Control* and one as *Autonomy*.

Fifteen of the 16 items loading on Factor III are classified as *Internal Control* and one as *Autonomy*.

From the foregoing it is clear that 24 of the 26 Autonomy items (according to the key) were correctly classified, 23 of the 28 External Control items were correctly classified, and 15 of the 26 Internal Control items were correctly classified. Eleven of the Internal Control items were misclassified as Autonomy. This is also evident in the matrix of intercorrelations of the factors (see Table 13): Factor I (Autonomy) correlates 0,496 with Factor III (Internal Control).

According to Table 10 the coefficients of skewness of the parcels loading on the first factor, range from –1,16 to –0,70, and those loading on the third factor, range from –0,54 to 0,01. The coefficients of skewness of the parcels loading on the second factor range from 0,01 to 0,77. The coefficients of skewness of the parcels loading on Factors I and III are thus essentially negative, whereas those of Factor II are positive.

TABLE 11
MATRIX OF INTERCORRELATIONS OF THE NEW VARIABLES ARRANGED ACCORDING TO THEIR DEGREE OF SKEWNESS

Variables	New var 2	New var 1	New var 3	New var 5	New var 10	New var 4	New var 7	New var 6	New var 9	New var 11
New var 2	1,000	0,577	0,546	0,390	0,330	0,383	0,326	0,392	0,500	0,360
New var 1	0,577	1,000	0,538	0,428	0,365	0,428	0,401	0,416	0,476	0,394
New var 3	0,546	0,538	1,000	0,397	0,363	0,394	0,372	0,397	0,464	0,386
New var 5	0,390	0,428	0,397	1,000	0,508	0,465	0,454	0,442	0,469	0,520
New var 10	0,330	0,365	0,363	0,508	1,000	0,435	0,458	0,471	0,443	0,526
New var 4	0,383	0,428	0,394	0,465	0,435	1,000	0,453	0,462	0,454	0,563
New var 7	0,326	0,401	0,372	0,454	0,458	0,453	1,000	0,491	0,442	0,545
New var 6	0,392	0,416	0,397	0,442	0,471	0,462	0,491	1,000	0,485	0,501
New var 9	0,500	0,476	0,464	0,469	0,443	0,454	0,442	0,485	1,000	0,486
New var 11	0,360	0,394	0,386	0,520	0,526	0,563	0,545	0,501	0,486	1,000
New var 8	0,423	0,461	0,427	0,461	0,392	0,445	0,429	0,493	0,480	0,488
New var 14	0,001	0,026	-0,055	0,095	0,140	0,108	0,115	0,113	0,046	0,143
New var 13	0,234	0,212	0,224	0,247	0,276	0,296	0,257	0,279	0,225	0,362
New var 12	0,174	0,162	0,209	0,284	0,286	0,264	0,256	0,190	0,230	0,306
New var 15	-0,080	-0,122	-0,137	-0,077	0,001	-0,084	-0,095	-0,063	-0,111	-0,075
New var 16	-0,077	-0,149	-0,192	-0,146	-0,090	-0,149	-0,202	-0,133	-0,164	-0,172
New var 17	-0,038	-0,116	-0,173	-0,040	0,048	-0,113	-0,090	-0,055	-0,066	-0,034
New var 19	-0,160	-0,247	-0,284	-0,206	-0,109	-0,216	-0,230	-0,197	-0,193	-0,159
New var 18	-0,130	-0,201	-0,230	-0,208	-0,091	-0,168	-0,163	-0,145	-0,140	-0,119
New var 20	-0,216	-0,269	-0,316	-0,190	-0,139	-0,236	-0,249	-0,167	-0,292	-0,216
Total	4,936	4,781	4,328	5,294	5,612	5,186	4,969	5,373	5,235	5,804

	New var 8	New var 14	New var 13	New var 12	New var 15	New var 16	New var 17	New var 19	New var 18	New var 20
	0,423	0,001	0,234	0,174	-0,080	-0,077	-0,038	-0,160	-0,130	-0,216
	0,461	0,026	0,212	0,162	-0,122	-0,149	-0,116	-0,247	-0,201	-0,269
	0,427	-0,055	0,224	0,209	-0,137	-0,192	-0,173	-0,284	-0,230	-0,316
	0,461	0,095	0,247	0,284	-0,077	-0,146	-0,040	-0,206	-0,208	-0,190
	0,392	0,140	0,276	0,286	0,001	-0,090	0,048	-0,109	-0,091	-0,139
	0,445	0,108	0,296	0,264	-0,084	-0,149	-0,113	-0,216	-0,168	-0,236
	0,429	0,115	0,257	0,256	-0,095	-0,202	-0,090	-0,230	-0,163	-0,249
	0,493	0,113	0,279	0,190	-0,063	-0,133	-0,055	-0,197	-0,145	-0,167
	0,480	0,046	0,225	0,230	-0,111	-0,164	-0,066	-0,193	-0,140	-0,292
	0,488	0,143	0,362	0,306	-0,075	-0,172	-0,034	-0,159	-0,119	-0,216
	1,000	0,045	0,249	0,206	-0,118	-0,142	-0,119	-0,228	-0,199	-0,226
	0,045	1,000	0,202	0,170	0,254	0,225	0,194	0,188	0,130	0,207
	0,249	0,202	1,000	0,195	0,066	0,042	0,081	0,014	0,014	-0,013
	0,206	0,170	0,195	1,000	0,004	-0,047	-0,071	-0,114	-0,147	-0,133
	-0,118	0,254	0,066	0,004	1,000	0,291	0,266	0,316	0,252	0,251
	-0,142	0,225	0,042	-0,047	0,291	1,000	0,374	0,375	0,316	0,415
	-0,119	0,194	0,081	-0,071	0,266	0,374	1,000	0,533	0,470	0,445
	-0,228	0,188	0,014	-0,114	0,316	0,375	0,533	1,000	0,555	0,381
	-0,199	0,130	0,014	-0,147	0,252	0,316	0,470	0,555	1,000	0,346
	-0,226	0,207	-0,013	-0,133	0,251	0,415	0,445	0,381	0,346	1,000
	4,965	3,348	4,462	3,423	1,737	1,375	2,494	1,017	1,143	0,382

TABLE 12
EIGENVALUES OF INTERCORRELATION MATRIX

Root	Eigenvalue
1	6,251
2	2,765
3	1,202
4	0,949
5	0,807
6	0,793
7	0,774
8	0,686
9	0,611
10	0,589
11	0,552
12	0,544
13	0,513
14	0,487
15	0,456
16	0,442
17	0,435
18	0,401
19	0,388
20	0,376
Trace	20,000

Table 13
Rotated factor matrix (direct oblimin)

VARIABLES	K	FACTOR I	FACTOR II	FACTOR III	h^2_j
New var 11: Items 5, 14, 44 and 74	4	**0,737**	-0,077	0,064	0,611
New var 10: Items 25, 29, 30 and 32	4	**0,616**	0,026	0,125	0,468
New var 7: Items 2, 40, 70 and 73	4	**0,599**	-0,155	0,091	0,470
New var 4: Items 8, 13, 28 and 62	4	**0,557**	-0,116	0,170	0,472
New var 5: Items 16, 22, 66 and 67	4	**0,548**	-0,085	0,195	0,471
New var 6: Items 24, 48, 54 and 69	4	**0,519**	-0,046	0,237	0,461
New var 12: Items 1, 9, 26 and 39	4	**0,430**	-0,088	-0,065	0,172
New var 13: Items 15, 47, 50 and 64	4	**0,416**	0,149	0,082	0,217
New var 8: Items 6, 46, 68 and 76	4	**0,406**	-0,103	0,326	0,444
New var 14: Items 3, 51, 71 and 77	4	**0,392**	0,310	-0,164	0,226
New var 17: Items 35, 36, 43 and 56	4	0,028	**0,747**	0,110	0,514
New var 19: Items 34, 41, 45 and 78	4	-0,100	**0,726**	0,017	0,540
New var 18: Items 12, 23, 33 and 79	4	-0,107	**0,643**	0,057	0,410
New var 16: Items 4, 20, 21 and 57	4	-0,066	**0,570**	0,025	0,325
New var 20: Items 11, 52, 53 and 58	4	-0,059	**0,551**	-0,138	0,391
New var 15: Items 38, 65, 72 and 80	4	0,086	**0,422**	-0,083	0,203
New var 2: Items 10, 18, 37 and 49	4	0,029	0,107	**0,795**	0,809
New var 1: Items 19, 42, 60 and 61	4	0,150	-0,039	**0,638**	0,544
New var 3: Items 31, 59, 63 and 75	4	0,136	-0,124	**0,591**	0,515
New var 9: Items 7, 17, 27 and 55	4	0,361	-0,046	**0,438**	0,498
Number of items per factor		40	24	16	

Intercorrelations of factors

VARIABLES	FACTOR I	FACTOR II	FACTOR III
FACTOR I	1,000	-0,086	**0,496**
FACTOR II	-0,086	1,000	**-0,339**
FACTOR III	**0,496**	**-0,339**	1,000

Note: Factor I = Autonomy
Factor II = External Control
Factor III = Internal Control

Table 14
Items correctly and incorrectly classified according to content

AUTONOMY			EXTERNAL CONTROL			INTERNAL CONTROL		
Q1	A		Q4	E		Q7	I	
Q2	A		Q11	E		Q10	I	
Q3	A		Q12	E		Q17	A	*
Q5	A		Q20	E		Q18	I	
Q6	I	*	Q21	E		Q19	I	
Q8	I	*	Q23	E		Q27	I	
Q9	E	**	Q33	E		Q31	I	
Q13	A		Q34	E		Q37	I	
Q14	A		Q35	E		Q42	I	
Q15	A		Q36	E		Q49	I	
Q16	I	*	Q38	E		Q55	I	
Q22	A		Q41	E		Q59	I	
Q24	A		Q43	E		Q60	I	
Q25	I	*	Q45	E		Q61	I	
Q26	I	*	Q52	E		Q63	I	
Q28	A		Q53	E		Q75	I	
Q29	A		Q56	E				
Q30	A		Q57	E				
Q32	I	*	Q58	E				
Q39	A		Q65	E				
Q40	I	*	Q72	A	*			
Q44	A		Q78	E				
Q46	A		Q79	E				
Q47	E	**	Q80	E				
Q48	I	*						
Q50	E	**						
Q51	E	**						
Q54	I	*						
Q62	A							
Q64	A							
Q66	A							
Q67	A							
Q68	A							
Q69	I	*						
Q70	A							
Q71	A							
Q73	A							
Q74	A							
Q76	I	*						
Q77	E	**						

Note: A = Autonomy; E = External Control; I = Internal Control
Items marked with * and ** have been misclassified according to the scoring key

Table 15
Matrix of Intercorrelations of the subtests of the locus of control inventory (1995)

Variable	Subtest 1	Subtest 2	Subtest 3	Subtest 4	Subtest 5	Subtest 6	Subtest 7	Subtest 8
Subtest 1	1,0000							
Subtest 2	-0,0660	1,0000						
Subtest 3	0,3673	-0,1870	1,0000					
Subtest 4	-0,1112	0,3609	-0,2250	1,0000				
Subtest 5	-0,1193	0,4261	-0,1823	0,3997	1,0000			
Subtest 6	0,5352	-0,1262	0,4588	-0,1844	-0,1430	1,0000		
Subtest 7	0,4443	-0,2841	0,2299	-0,3492	-0,3382	0,2890	1,0000	
Subtest 8	0,2911	-0,1325	0,3450	-0,1474	-0,1492	0,3189	0,2669	1,0000
Subtest 9	0,5430	-0,1719	0,3074	-0,1914	-0,2084	0,3998	0,4747	0,3398
Subtest 10	0,3192	0,0411	0,3547	-0,0128	0,0477	0,3824	0,0900	0,1494
Subtest 11	0,3171	-0,0056	0,4365	-0,0927	0,0012	0,3209	0,1342	0,1703
Subtest 12	0,3335	-0,0827	0,4752	-0,0827	-0,0803	0,3639	0,1527	0,3046
Subtest 13	-0,1734	0,1992	-0,0888	0,2190	0,2050	-0,1807	-0,2647	0,0033
Subtest 14	-0,2024	0,3260	-0,1508	0,3008	0,3533	-0,2132	-0,3866	-0,1948
Subtest 15	-0,1889	0,1709	-0,1242	0,2012	0,3143	-0,1844	-0,2833	-0,2028
Subtest 16	0,0403	0,1888	0,1442	0,1621	0,3073	0,0453	-0,1050	0,0330
Subtest 17	0,5388	-0,1952	0,4480	-0,2378	-0,1795	0,4685	-0,4335	0,3277

Subtest 9	Subtest 10	Subtest 11	Subtest 12	Subtest 13	Subtest 14	Subtest 15	Subtest 16	Subtest 17
1,0000								
0,2278	1,0000							
0,2094	0,3938	1,0000						
0,2518	0,2506	0,2515	1,0000					
-0,1470	-0,0074	-0,0225	-0,0553	1,0000				
-0,2475	-0,0021	-0,0424	-0,0922	0,2427	1,0000			
-0,2405	-0,0590	-0,0711	-0,1527	0,1579	0,2135	1,0000		
-0,0269	0,1719	0,1824	0,1467	0,1067	0,1589	0,1321	1,0000	
0,5171	0,3308	0,3401	0,3231	-0,2012	-0,2200	-0,2156	0,0121	1,0000

Table 16
Eigenvalues of unreduced intercorrelation matrix (17 x 17)

Root	Eigenvalue
1	4,727540
2	2,373090
3	1,103630
4	1,002740
5	0,861917
6	0,837644
7	0,773758
8	0,695669
9	0,650908
10	0,620371
11	0,606981
12	0,532800
13	0,521134
14	0,466443
15	0,450356
16	0,411529
17	0,363493
Trace	17,000000

Objective 3: Overcoming the effects of differential skewness of test items in scale construction

As the procedure that was followed is fully described in the method section, only the essential results are given here.

The items of the LCI were intercorrelated, and the *eigenvalues* of the intercorrelation matrix were calculated. Nineteen of the eigenvalues were greater than unity, accordingly 19 factors were extracted and rotated to simple structure by means of a Varimax rotation.

Two of the factors had *one* loading each and were discarded. Next, 17 *subscores* were formed by adding all the items with substantial loadings on a factor, together. The 17 subscores were then intercorrelated. The matrix of intercorrelations is given in Table 15.

From Table 15 it is clear that the correlations of the subscores with one another vary from moderate to low and from positive to negative, suggesting more than one factor.

Next, the eigenvalues of the intercorrelation matrix were calculated. The obtained eigenvalues are given in Table 16.

Four of the eigenvalues were greater than unity, suggesting four factors (Kaiser, 1961).

Accordingly four factors were extracted and rotated to simple structure by means of a Direct Oblimin rotation. The rotated factor matrix is given in Table 17.

From an inspection of Table 17 it is clear that the first three factors are *well determined* with four or more high loadings. However, the fourth factor had only one high loading. A three-factor-solution was therefore tried. The obtained factor matrix is given in Table 18.

From Table 18 it is clear that all three factors are *well determined* with *four or more* high loadings. From the intercorrelations of the factors it is clear that *External Control* and *Internal Control* are essentially *uncorrelated*. External Control is moderately negatively correlated with Autonomy, and Internal Control is moderately *positively* correlated with Autonomy.

Table 17
Rotated factor marix (direct oblimin)

VARIABLES	K	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h_j^2
Subtest 1: Items 2, 3, 5, 14, 15, 22, 24, 28, 29, 62, 64, 67 and 70	13	0,738	0,131	+0,192	0,016	0,634
Subtest 2: Items 12, 34, 35, 36, 41 and 79	6	0,081	0,627	-0,039	-0,050	0,365
Subtest 3: Items 10, 42, 49, 61, 63 and 75	6	-0,090	-0,275	+0,694	0,244	0,656
Subtest 4: Items 20, 43, 52, 53, 56 and 78	6	0,027	0,607	-0,123	0,024	0,360
Subtest 5: Items 9, 51, 57, 58, 65, 77 and 80	7	0,022	**0,698**	+0,050	-0,080	0,499
Subtest 6: Items 6, 7, 16, 25, 37, 59 and 69	7	0,362	-0,080	+0,403	0,067	0,464
Subtest 7: Items 1, 17, 39, 44, 71 and 72	6	0,546	0,296	-0,070	0,020	0,504
Subtest 8: Items 8, 40 and 54	3	0,228	-0,040	+0,040	0,516	0,405
Subtest 9: Items 30, 46, 73 and 74	4	**0,669**	-0,001	+0,004	0,155	0,521
Subtest 10: Items 26, 27, 31 and 32	4	0,155	0,108	+0,537	-0,060	0,357
Subtest 11: Items 18 and 19	2	0,059	-0,010	+0,595	-0,050	0,368
Subtest 12: Items 48, 55, 60 and 76	4	0,073	-0,040	+0,374	0,308	0,346
Subtest 13: Items 21, 23 and 33	3	-0,185	0,315	-0,045	0,201	0,191
Subtest 14: Items 38 and 45	2	-0,194	0,442	+0,050	-0,040	0,308
Subtest 15: Items 4 and 11	2	-0,203	0,262	+0,050	-0,130	0,183
Subtest 16: Items 47 and 50	2	-0,050	0,331	+0,295	0,054	0,211
Subtest 17: Items 13, 66 and 68	3	0,494	-0,109	+0,310	0,045	0,532
Number of items per factor		26	28	23	3	

Note: Factor 3 has been reflected

Table 18
Rotated factor matrix (direct oblimin)

VARIABLES	K	FACTOR I	FACTOR II	FACTOR III	h_j^2
Subtest 1: Items 2, 3, 5, 14, 15, 22, 24, 28, 29, 62, 64, 67 en 70	13	0,146	0,160	0,754	0,634
Subtest 2: Items 12, 34, 35, 36, 41 en 79	6	0,630	-0,083	0,082	0,366
Subtest 3: Items 10, 42, 49, 61, 63 en 75	6	-0,266	0,828	-0,105	0,678
Subtest 4: Items 20, 43, 52, 53, 56 en 78	6	0,579	-0,119	0,016	0,343
Subtest 5: Items 9, 51, 57, 58, 65, 77 en 80	7	0,717	-0,017	0,024	0,502
Subtest 6: Items 6, 7, 16, 25, 37, 59 en 69	7	-0,052	0,411	0,374	0,462
Subtest 7: Items 1, 17, 39, 44, 71 en 72	6	0,304	0,070	0,557	0,506
Subtest 8: Items 8, 40 en 54	3	-0,135	0,283	0,211	0,219
Subtest 9: Items 30, 46, 73 en 74	4	-0,037	0,065	0,665	0,505
Subtest 10: Items 26, 27, 31 en 32	4	0,169	0,464	0,173	0,318
Subtest 11: Items 18 en 19	2	0,056	0,529	0,083	0,324
Subtest 12: Items 48, 55, 60 en 76	4	-0,064	0,514	0,074	0,312
Subtest 13: Items 21, 23 en 33	3	0,262	0,057	-0,193	0,139
Subtest 14: Items 38 en 45	2	0,451	0,027	-0,198	0,309
Subtest 15: Items 4 en 11	2	0,291	-0,018	-0,203	0,176
Subtest 16: Items 47 en 50	2	0,346	0,307	-0,048	0,213
Subtest 17: Items 13, 66 en 68	3	-0,087	0,306	0,507	0,530
Number of items per factor		28	26	26	80

Intercorrelations of factors

VARIABLES	FACTOR I	FACTOR II	FACTOR III
FACTOR I	1,000		
FACTOR II	-0,018	1,000	
FACTOR III	-0,393	0,438	1,000

Note: N = 1662
Factor I = External Control
Factor II = Internal Control
Factor III = Autonomy

Next, separate scales were formed, corresponding to each of the factors, and subjected to item analysis.

The item statistics in respect of Scale I (External Control) are given in Table 19.

From Table 19 it is clear that the *item-total* correlations range from 0,341 to 0,611, with a mean of 0,459 and a standard deviation of 0,080. The scale is therefore internally highly consistent. Items 23, 33 and 50 were rejected, because their indices of reliability were too low. The reliability of the scale according to Cronbach's coefficient alpha is 0,841.

The item statistics in respect of Scale II (Internal Control) are given in Table 20.

According to Table 20 the *item-total* correlations range from 0,301 to 0,585, with a mean of 0,455 and a standard deviation of 0,069. The scale is therefore internally highly consistent. No items were rejected. The reliability of the scale according to Cronbach's coefficient alpha is 0,832.

The item statistics in respect of Scale III (Autonomy) are given in Table 21.

From Table 21 it is clear that the item-total correlations range from 0,370 to 0,575, with a mean of 0,488 and a standard deviation of 0,074. Therefore, the scale is internally highly consistent. No items were rejected. The reliability of the scale according to Cronbach's coefficient alpha is 0,866.

Table 19
Item statistics in respect of scale I of the LCI: External control

	DESCRIPTION OF ITEM	N	MEAN OF ITEM ($\bar{X}_g$)	STANDARD DEVIATION OF ITEM (s_g)	ITEM-TEST CORRELATION (r_{gx})	INDEX OF RELIABILITY OF ITEM ($r_{gx}s_g$)
Q4	Convinced that a person without money will get nowhere, no matter how hard he/she works	1662	3,620	1,835	0,341	0,625
Q9	People obtain good positions simply because they know the right people	1662	5,039	1,291	0,392	0,506
Q11	Convinced that if once he/she has failed in something it is virtually impossible to achieve in it again	1662	2,471	1,392	0,379	0,528
Q12	Convinced that he/she is subject to the whims of fate	1662	2,788	1,518	0,573	0,869
Q20	Failed because other people interfered in his/her business	1662	3,397	1,367	0,409	0,560
Q21	Dependent on the advice and cues of others in order to produce quality work	1662	3,426	1,325	0,398	0,527
Q23	Will readily support a group decision even if he/she disagrees with it	1662	2,978	1,444	****	****
Q33	Will readily accept responsibility for errors in work situation even if he/she is innocent	1662	2,801	1,460	****	****
Q34	Lady Luck plays an important role in his/her life	1662	2,644	1,518	0,471	0,715
Q35	Strongly believes in fatalism	1662	2,867	1,525	0,472	0,720
Q36	Life influenced to a great extent by coincidences	1662	3,195	1,455	0,529	0,770
Q38	Other people are responsible for his/her wellbeing	1662	3,581	1,437	0,360	0,518
Q41	Convinced that failures in life could be attributed to fate	1662	2,549	1,441	0,570	0,821
Q43	Present achievements adversely affected by negative experiences in past	1662	3,127	1,626	0,468	0,761
Q45	Other people are in charge of his/her life and determine outcome of issues	1662	2,505	1,359	0,536	0,729
Q47	Convinced that a person cannot achieve without the right opportunities	1662	4,619	1,445	0,366	0,529
Q50	Convinced that success depends mainly upon equal opportunities in life	1662	4,757	1,376	****	****
Q51	Advancement in life is determined by your superiors	1662	3,652	1,375	0,496	0,683
Q52	Parents/guardians negatively influenced his/her achievement at school	1662	2,142	1,467	0,391	0,574
Q53	Present achievement negatively influenced by people who are not favourably disposed towards him/her	1662	2,172	1,304	0,485	0,632
Q56	His/her progress in the past thwarted by people that were hostile towards him/her	1662	3,035	1,498	0,462	0,692
Q57	Convinced that only people who are at the right place at the right time get promoted	1662	3,282	1,465	0,569	0,833
Q58	Only people who belong to the right political party have a chance in life	1662	2,342	1,564	0,484	0,757
Q65	Convinced that clique formation is the most important determinant of social acceptance	1662	3,710	1,676	0,398	0,668
Q77	Convinced that promotion in the new South Africa will depend largely on skin-colour	1662	4,306	1,784	0,423	0,755
Q78	Convinced that it is impossible to rise above your own environment	1662	2,936	1,690	0,344	0,582
Q79	Convinced that his/her fate is determined by coincidental events over which he/she has no control	1662	2,690	1,463	0,611	0,894
Q80	Convinced that his/her advancement in life will be determined by certain influential people	1662	3,634	1,541	0,540	0,831

Means and standard deviations of item statistics
(only in respect of items included in test score)

	$\bar{X}_g$	s_g	r_{gx}	$r_{gx}s_g$
Mean	3,189	1,494	0,459	0,683
SD	0,736	0,141	0,080	0,120

Cronbach alpha = 0,841
Mean of test = 79,730
Standard deviation = 17,079
Number of items = 25

Table 20
Item statistics in respect of scale II of the LCI: internal control

	DESCRIPTION OF ITEM	N	MEAN OF ITEM ($\overline{X}_g$)	STANDARD DEVIATION OF ITEM (s_g)	ITEM-TEST CORRELATION (r_{gx})	INDEX OF RELIABILITY OF ITEM ($r_{gx}s_g$)
A6	Convinced that personal insight is a pre-requisite for good interpersonal relationships	1662	5,686	1,096	0,458	0,502
A7	The structure and routine of a person's work should be determined by himself/herself	1662	5,664	1,045	0,465	0,486
A8	Readily accepts responsibility for mistakes that appear in his/her work	1662	5,607	1,149	0,386	0,444
A10	Convinced that success is mainly related to a person's ability and dedication	1662	6,126	0,930	0,475	0,442
A16	Decides on matters himself/herself, rather than waiting for others to take decisions on his/her behalf	1662	5,180	1,619	0,343	0,556
A18	Recognition encourages him/her to perform even better	1662	6,063	1,018	0,461	0,470
A19	Success encourages him/her to work harder and achieve greater heights	1662	6,329	0,854	0,525	0,448
A25	Prefers to follow his/her own mind, rather than following someone else's instructions	1662	5,274	1,303	0,389	0,507
A26	Insists on recognition of his/her own individual achievements	1662	4,655	1,571	0,301	0,473
A27	Takes responsibility for his/her own intellectual development	1662	5,623	1,088	0,483	0,526
A31	Important for him/her to receive feedback on tasks which he/she has performed	1662	5,821	1,122	0,471	0,528
A32	Convinced that reward for achievement is earned	1662	5,252	1,234	0,454	0,560
A37	Convinced that achievement of personal objectives depends on oneself	1662	5,954	1,043	0,539	0,562
A40	Readily accepts responsibility for his/her own poor performance	1662	5,306	1,329	0,407	0,541
A42	Convinced that the respect one receives is directly related to one's behaviour	1662	6,100	1,059	0,538	0,570
A48	Agrees that failure in life can be attributed to a lack of dedication	1662	5,354	1,303	0,392	0,511
A49	Convinced that success depends mainly on hard work	1662	6,220	0,899	0,515	0,463
A54	Takes personal responsibility for things that go wrong in his/her life	1662	5,427	1,177	0,427	0,503
A55	Convinced that the outcome of matters is determined by his/her own inputs	1662	5,589	1,000	0,585	0,585
*A59	Believes that his/her own input bears no relation to the outcome of matters	1662	5,744	1,351	0,390	0,527
A60	Convinced that achievement depends upon his/her utilising his/her God-given talents to the full	1662	6,133	1,120	0,446	0,499
A61	Convinced that the achievements he/she has obtained were deserved	1662	5,968	1,131	0,490	0,554
A63	Convinced that promotions are earned through hard work and perseverance	1662	6,194	0,900	0,483	0,434
A69	Convinced that he/she is sufficiently qualified for the work he/she is doing	1662	5,742	1,107	0,493	0,546
A75	Convinced that the achievements he/she has obtained are the results of hard work and dedication	1662	5,969	1,087	0,552	0,600
A76	Convinced that failures in life are due to a lack of perseverance	1662	5,023	1,519	0,356	0,541

Means and standard deviations of item statistics
(only in respect of items included in test score)

	$\overline{X}_g$	s_g	r_{gx}	$r_{gx}s_g$
Mean	5,692	1,156	0,455	0,515
SD	0,418	0,201	0,069	0,047

Cronbach alpha = 0,832
Mean of test = 148,001
Standard deviation = 13,359
Number of items = 26

Table 21
Item statistics in respect of scale III of the LCI: autonomy

	DESCRIPTION OF ITEM	N	MEAN OF ITEM ($\bar{X}_g$)	STANDARD DEVIATION OF ITEM (s_g)	ITEM-TEST CORRELATION (r_{gx})	INDEX OF RELIABILITY OF ITEM ($r_{gx}s_g$)
*Q1	Doubts his/her own capabilities when his/her work is being criticised	1662	4,796	1,372	0,468	0,642
Q2	Geared towards ensuring that his/her case triumphs during a conflict situation	1662	5,494	1,277	0,429	0,547
Q3	Would readily take risks	1662	4,500	1,363	0,370	0,505
Q5	Can readily convince someone else of his/her viewpoint	1662	5,256	1,017	0,559	0,569
Q13	Convinced that he/she will succeed when undertaking important tasks	1662	5,797	0,951	0,566	0,539
Q14	Makes things happen through his/her own input, rather than waiting for things to happen	1662	5,201	1,165	0,575	0,670
*Q15	Waits for other people to take charge, rather than taking charge	1662	4,691	1,503	0,543	0,816
Q17	Failure spurs him/her on to improve his/her performance	1662	5,315	1,288	0,392	0,505
Q22	Likes taking decisions himself/herself	1662	5,804	1,102	0,574	0,633
Q24	Would readily air his/her views when they differ from someone else's	1662	5,440	1,249	0,549	0,686
Q28	Likes occupying a leadership position	1662	5,463	1,366	0,533	0,728
Q29	Would stick to his/her viewpoint when someone for whom he/she has great respect disagrees with him/her	1662	5,288	1,211	0,494	0,598
Q30	Likes solving complex problems	1662	5,206	1,325	0,518	0,686
*Q39	Feels that he/she has no control over his/her own circumstances	1662	4,779	1,400	0,407	0,570
Q44	Often achieves set objectives, irrespective of the conditions	1662	5,187	1,023	0,533	0,545
Q46	Convinced that he/she can solve most problems, irrespective of the conditions	1662	5,400	1,113	0,504	0,561
Q62	Can predict the outcome of an examination he/she has just written	1662	5,623	1,205	0,335	0,403
Q64	Finds it easy to satisfy choosy people	1662	4,345	1,296	0,381	0,494
Q66	Convinced that he/she possesses the ability to produce work of the highest quality	1662	6,090	0,902	0,517	0,466
Q67	Would strongly defend his/her actions if the appropriateness thereof were to be questioned by others	1662	5,764	1,055	0,517	0,545
Q68	Convinced that he/she is sufficiently qualified for the work that he/she is doing	1662	5,739	1,007	0,546	0,550
Q70	Prefers challenging work to routine work	1662	5,649	1,217	0,518	0,630
*Q71	Subsequently doubts the correctness of the decisions he/she has taken	1662	4,430	1,292	0,433	0,559
*Q72	Dependent on the support and goodwill of others in the execution of tasks	1662	4,294	1,337	0,365	0,488
*Q73	Would readily quit if he/she is battling with a complex problem	1662	5,199	1,282	0,483	0,619
Q74	Often takes the initiative in finding solutions for troublesome problems	1662	5,210	1,129	0,575	0,649

Means and standard deviations of item statistics
(only in respect of items included in test score)

	$\bar{X}_g$	s_g	r_{gx}	$r_{gx}s_g$
Mean	5,229	1,210	0,488	0,585
SD	0,489	0,153	0,074	0,090

Cronbach alpha = 0,866
Mean of test = 135,958
Standard deviation = 15,189
Number of items = 26

DISCUSSION

From the findings of the study it is clear that the degree of skewness (marginal splits) of *binary* items places an *upper limit* on the correlations between the items. Furthermore, intercorrelation matrices based on such items, arranged according to their degree of skewness, have the typical structure of a simplex, quasi-simplex or pseudosimplex. Factoring such matrices result in factors of skewness, regardless of the contents of the items.

As far as the first objective of the study is concerned, it was found that similar *gradients* exist in respect of the correlations in an intercorrelation matrix based on *continuous* variables, arranged in terms of skewness.

To achieve the second objective, the items of the LCI were grouped into parcels of four items each. The parcels were intercorrelated and subjected to factor analysis. Three factors were obtained. These factors were also largely factors of skewness. Content also played a role in so far as the contents of the items were associated with their degree of skewness. The items loading on Factors I and III (Autonomy and Internal Control) are essentially *negatively* skewed, whereas those loading on Factor II (External Control) are *positively* skewed.

As far as the third and major objective of the study is concerned, the procedure that was followed, yielded three factors that were well determined. The corresponding scales that were produced are internally highly consistent, with reliabilities that range from 0,832 to 0,866.

Autonomy is *positively* correlated with Internal Control ($r = 0{,}438$; $p < 0{,}001$). There is thus approximately 19% *common variance* between the two constructs. However, the reliable variance of Autonomy is 87% and that of Internal Control 83%. The specific variances of the two scales therefore vary from 64%to 68%.

Autonomy and External Control are negatively correlated ($r = -0{,}393$; $p < 0{,}001$). There is thus approximately 15% common variance between the two constructs. The reliable variance of External Control is 84%. The specific variances of the two constructs (scales) therefore vary from 69% to 72%.

Internal Control and External Control are essentially uncorrelated ($r = -0{,}018$; $p > 0{,}05$). This is in keeping with Social Learning Theory and Attribution Theory: The causes of human behaviour can be divided into two broad classes, namely those that pertain to one's natural disposition and those that pertain to the external world (Roediger III et al., 1991). It is therefore not surprising that items from these two domains are essentially *independent* of one another.

In the procedure followed in the present study, the items were grouped according to the factors they loaded on. Each subscore was therefore *internally consistent*. The fact that a Varimax rotation was used, kept the factors relatively independent of one another, and simplified the procedure. In intercorrelating the items of the LCI both content and degree of skewness of the items must have played a role. However, the degree of skewness of the subtests are considerably smaller than those of the single items.

Forming parcels on a priori theoretical grounds does not guarantee internal consistency within parcels or eliminate the effects of differential skewness completely. For this, a *new measure* of *association* that is *independent* of skewness, is required.

The procedure described, has been used at the Rand Afrikaans University since 1992, and has consistently produced scales of high reliability and validity.

ACKNOWLEDGEMENTS

I hereby wish to thank all the members of the Statistical Consultation Service of the Rand Afrikaans University for all the hours of computational work done for me. I value it very highly.

A special word of thanks to Annetjie Boshoff for the typing of the manuscript. Nobody can type tables better than her.

REFERENCES

Baron, R.A., Byrne, D. & Kantowitz, B.H. (1980). *Understanding behavior* (2nd ed.) New York: Holt, Rinehart and Winston.

Bohrnstedt, G.W. & Knoke, D. (1988). *Statistics for social data analysis* (2nd ed.) Itasca, Ill.: F.E. Peacock Publishers, Inc..

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Ferguson, G.A. (1941). The factorial interpretation of test difficulty. *Psychometrika*, *6*, 323-329.

Gorsuch, R.L. (1974). *Factor analysis*. Philadelphia: W.B. Saunders Company.

Guilford, J.P. (1950). *Fundamental statistics in psychology and education* (2nd ed.) New York: McGraw-Hill.

Guttman, L.A. (1954). A new approach to factor analysis: the radex. In: Lazarsfeld, P.F. (Ed.): *Mathematical thinking in the social sciences*. Glencoe, Ill: The Free Press.

Guttman, L.A. (1955). A generalized simplex for factor analysis. *Psychometrika*, *20*, 173-192

Guttman, L.A. (1957). Empirical verification of the radex structure of mental abilities and personality traits. *Educational and Psychological Measurement*, *17*, 391-407.

Horst, P. (1953). Correcting the Kuder-Richardson reliability for dispersion of item difficulties. *Psychological Bulletin*, *50*, 371-374.

Horst, P. (1965). *Factor analysis of data matrices*. New York: Holt, Rinehart and Winston.

Jöreskog, K.G. (1970). Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology*, 23, 121-145.

Jöreskog, K.G. & Sorbom, D. (1982). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares*. Uppsala, Sweden: University of Uppsala.

Kaiser, H.F. (1961). A note on Guttman's lower bound for the number of common factors. *British Journal of Statistical Psychology*, *14* (1), 1.

Kuder, G.F. & Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*, 151-160.

Magnusson, D. (1967). *Test theory*. Reading Mass.: Addison-Wesley.

Mischel, W. (1979). On the interface of cognition and personality: Beyond the person-situation debate. *American Psychologist*, *34*, 740-754.

Roediger III, H.L., Capaldi, E.D., Paris, S.G. & Polivy, J. (1991). *Psychology* (3rd ed.) New York: Harper Collins Publishers.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, *80*, No 1 (Whole No. 609).

Schepers, J.M. (1962). A components analysis of a complex psychomotor learning task. *Psychologia Africana*, *9*, 294-329.

Schepers, J.M. (1992). *Toetskonstruksie: Teorie en praktyk*. Johannesburg: RAU-Drukpers.

Schepers, J.M. (1995). *The Locus of Control Inventory* (revised edition). Johannesburg.

Tucker, L.R. (1949). A note on the estimation of test reliability by the Kuder-Richardson Formula (20). *Psychometrika*, *14*, 117-119.

Wolman, B.B. (1973). *Dictionary of behavioral science*. New York: Van Nostrand Reinhold Company.

APPENDIX 1

Cronbach's coefficient alpha and Kuder-Richardson Formula 20 (KR20), for continuous data, are formally the same, and can be written as follows:

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{\sum_{g=1}^{K}\sigma_g^2}{\sigma_x^2}\right], \text{ where} \tag{1}$$

K = number of test items

$\sum \sigma_g^2$ = sum of item variances

σ_g^2 = test variance

For **binary** data KR_{20} can be written as follows:

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{\sum_{g=1}^{K}(p_g - p_g^2)}{\sigma_x^2}\right], \text{ where} \tag{2}$$

K = number of test items

p_g = proportion of subjects endorsing item g according to the key

The variance of the p-values can be written as

$$\sigma_p^2 = \frac{\sum_{g=1}^{K}(p_g - \bar{p})^2}{K}, \tag{3}$$

$$= \frac{\sum p_g^2}{K} - \bar{p}^2$$

From this it follows that

$$\sum_{g=1}^{K} p_g^2 = K(\sigma_p^2 + \bar{p}^2)$$

Substitution for $\sum_{g=1}^{K} p_g^2$ in (2) gives

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{\sum_{g=1}^{K} p_g - K(\sigma_p^2 + \bar{p}^2)}{\sigma_x^2}\right], \tag{4}$$

$$= \frac{K}{K-1}\left[1 - \frac{\mu - K(\sigma_p^2 + \frac{\mu^2}{K^2})}{\sigma_x^2}\right], \tag{5}$$

Because $\sum_{g=1}^{K} p_g^2 = \mu$ (mean of test)

and $\bar{p} = \frac{\mu}{K}$

$$\therefore KR_{20} = \frac{K}{K-1}\left[1 - \frac{\mu - \frac{\mu^2}{K} - K\sigma_p^2}{\sigma_x^2}\right] \tag{6}$$

Formula 6 applies only to binary data.

For continuous data Kuder-Richardson Formula 20 and Cronbach's coefficient alpha can be written as follows:

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{\sum \sigma_g^2}{\sigma_x^2}\right], \text{ where} \tag{7}$$

K = number of test items

$\sum \sigma_g^2$ = sum of item variances

σ_x^2 = test variance

Formula 7 can be transformed as follows:

$$KR_{20} = \frac{K}{K-1}\left[1 - \frac{1}{K} - \frac{\sum_{g=1}^{K}\sigma_g^2}{\sigma_x^2} + \frac{1}{K}\right], \tag{8}$$

$$= \left[1 - \frac{(K\sum_{g=1}^{K}\sigma_g^2 - \sigma_x^2)}{(K-1)\sigma_x^2}\right], \tag{9}$$

$$= \left[1 - \frac{K\sum_{g=1}^{K}\sigma_g^2 - \sigma_x^2}{(K-1)\sigma_x^2}\right], \tag{10}$$

The total **test variance** can never exceed $K\sum_{g=1}^{K}\sigma_g^2$ therefore the greater the test variance, the higher the reliability (Schepers, 1992, p. 33).

APPENDIX 2

PHI COEFFICIENT IN RESPECT OF BINARY TEST ITEMS

Item 1

Item 2	+	–	**Total**
–	A 30 (40)	B 20 (10)	A+B 50 (50) q_k = 0,5
+	C 10 (0)	D 40 (50)	C+D 50 (50) p_k = 0,5
Total	A+C 40 (40) p_g = 0,4	B+D 60 (60) q_g = 0,6	N 100 (100)

$$r_{phi} = \frac{p_{gk} - p_g p_k}{\sqrt{p_g q_g \, p_k q_k}}$$

$$= \frac{p_{gk} - 0,4 \times 0,5}{\sqrt{0,4 \times 0,6 \times 0,5 \times 05}} = \frac{0,10 - 0,20}{0,244948974}$$

$$= -0,40824829$$

$$\phi = \frac{BC - AD}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

$$= \frac{200 - 1200}{2449,489743}$$

$$= -0,40824829 \Rightarrow$$

Maximum value of ϕ
Set $C = 0$*, then*

$$\phi = \frac{0 - 2000}{2449,489743}$$

$$= -0,81649658 \Rightarrow$$

$$_{phi} = \frac{p_{gk} - p_g p_k}{\sqrt{p_g q_g p_k q_k}}$$

$$= \frac{0,4 - 0,4 \times 0,5}{\sqrt{0,4 \times 0,6 \times 0,5 \times 0,5}}$$

$$= \frac{0,20}{0,244948974}$$

$$= 0,81649658 \Rightarrow$$

$$\phi = \frac{p_{gk} - p_g p_k}{\sqrt{p_g q_g p_k q_k}}$$

To obtain $\phi_{\max}$*, set* $p_{gk} = p_g$*, where* $p_g \leq p_k$*,*

$$\phi_{\max} = \frac{p_g - p_g p_k}{\sqrt{p_g q_g p_k q_k}}$$

$$= \frac{p_g (1 - p_k)}{\sqrt{p_g q_g p_k q_k}}$$

$$= \frac{p_g q_k}{\sqrt{p_g q_g p_k q_k}}$$

$$\phi^2_{\max} = \frac{p_g^2 q_k^2}{p_g q_g p_k q_k}$$

$$\phi_{\max} = \sqrt{\frac{p_g q_k}{q_g p_k}}$$

$$= \sqrt{\left(\frac{p_g}{q_g}\right)}\left(\frac{q_k}{p_k}\right)$$

Note that the maximum value of ϕ is a direct function of the **marginal splits** $(p_g | q_g \text{ and } q_k | p_k)$.

Example

$$\phi_{\max} = \sqrt{\left(\frac{0,4}{0,6}\right)\left(\frac{0,5}{0,5}\right)} = \sqrt{\frac{0,20}{0,30}} = 0,81649658$$