PSYCHOMETRIKA—VOL. 39, NO. 1
MARCH, 1974

AN INDEX OF FACTORIAL SIMPLICITY*

HENRY F. KAISER

UNIVERSITY OF CALIFORNIA, BERKELEY

AND

UNITED STATES COAST GUARD ACADEMY

An index of factorial simplicity, employing the quartimax transformational criteria of Carroll, Wrigley and Neuhaus, and Saunders, is developed. This index is both for each row separately and for a factor pattern matrix as a whole. The index varies between zero and one. The problem of calibrating the index is discussed.

After a factor analysis has been completed, it is of interest to assess how good the solution is, in the sense of how simple—and thus how interpretable—the final factor pattern matrix is. The ideal solution, most investigators would agree, is one that is unifactorial, *i.e.*, a solution for which each row of the pattern matrix has one, and only one, non-zero loading. In this paper we propose an index, for each row of the factor matrix separately and for the matrix as a whole, which measures the tendency towards unifactoriality for a given row and the tendency toward unifactoriality for the entire factor pattern matrix. For this purpose, we turn to the quartimax criteria for analytic transformation of Carroll [1953], Wrigley and Neuhaus [1954], and Saunders [1953], and measure the value of their criteria for the data at hand relative to the optimum value of their criteria. Interestingly, it turns out that each of their three rationales leads to the same index of factorial simplicity when we scale the index to lie between zero and one. (It should be pointed out that, while we develop our index from the quartimax viewpoint, our results are applicable to any factor pattern matrix.)

First we consider Carroll's [1953] criterion. For row j of the pattern matrix V, he proposes that

$$C_j = \sum_{s \neq t} \sum v_{js}^2 v_{jt}^2 \tag{1}$$

should be a minimum. Clearly the minimum value that C_j can reach is zero, and this occurs when only one of the q (the number of factors) elements in row j is non-zero. The worst possible value of C_j, *i.e.*, when C_j is a maximum,

* The research reported in this paper was supported in part by the Office of Computing Activities, National Science Foundation.

occurs when all q of the v_{is} in row j are equal in absolute value. It is seen that this is given by

$$\text{(2)} \qquad \max C_j = q(q-1)(\sum_s v_{js}^2/q)^2,$$

$$\text{(3)} \qquad = (q-1)(\sum_s v_{js}^2)^2/q.$$

Let us take as our squared index of factorial simplicity, for row j of V, the quantity

$$\text{(4)} \qquad (IFS(J))^2 = 1 - \frac{C_j}{\max C_j}$$

$$\text{(5)} \qquad = 1 - \frac{\sum\sum_{s \neq t} v_{js}^2 v_{jt}^2}{(q-1)(\sum_s v_{js}^2)/q}$$

We define our $IFS(J)$ in this way so that it can vary from zero to one, yielding a value of zero when all v_{js} are equal in absolute value—the worst possible situation—and reaching a value of one only when all v_{js} but one are zero—the best possible situation. (We discuss why we take the square root of (5) for the actual $(IFS(J))$ below under "the problem of calibration.") Noting that

$$\text{(6)} \qquad \sum\sum_{s \neq t} v_{js}^2 v_{jt}^2 = (\sum_s v_{js}^2)^2 - \sum_s v_{js}^4,$$

and after a little algebra, we find

$$\text{(7)} \qquad (IFS(J))^2 = \frac{q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2}{(q-1)(\sum_s v_{js}^2)^2}.$$

From Carroll's viewpoint, his overall criterion for the entire pattern matrix is

$$\text{(8)} \qquad C = \sum_j C_j = \sum_j \sum\sum_{s \neq t} v_{js}^2 v_{jt}^2.$$

If we define our overall squared index of factorial simplicity as

$$\text{(9)} \qquad (IFS)^2 = 1 - \frac{C}{\max C},$$

we readily find that

$$\text{(10)} \qquad (IFS)^2 = \frac{\sum_j [q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2]}{\sum_j [(q-1)(\sum_s v_{js}^2)^2]}$$

which again can vary from zero—worst possible—to one—best possible. The latter occurs only when the entire matrix is unifactorial.

From Wrigley and Neuhaus' [1954] point of view a row of V is ideal when the variance of the squared loadings is a maximum, *i.e.*, when

$$W_j = [q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2]/q^2 \tag{11}$$

is a maximum. The maximum value that W_j can attain is

$$\max W_j = [q(\sum_s v_{js}^2)^2 - (\sum_s v_{js}^2)^2]/q^2. \tag{12}$$

From this viewpoint let us take as our squared index of factorial simplicity

$$(IFS(J))^2 = \frac{W_j}{\max W_j}, \tag{13}$$

$$= \frac{q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2}{(q-1)(\sum_s v_{js}^2)^2}, \tag{14}$$

the same index we proposed when looking at the problem from Carroll's viewpoint. Overall—for the entire matrix—we define

$$(IFS)^2 = \frac{\sum_j W_j}{\sum_j \max W_j} \tag{15}$$

$$= \frac{\sum_j [q \sum_s v_{js}^4 - (\sum_s v_{js}^2)^2]}{\sum_j [(q-1)(\sum_j v_{js}^2)^2]}, \tag{16}$$

as before.

Finally, Saunders' [1953] version of the quartimax criterion proposes that, for row j,

$$S_j^* = \frac{\sum_s v_{js}^4}{(\sum_s v_{js}^2)^2} \tag{17}$$

should be a maximum. The maximum value of S_j^* occurs, as might be expected, when one and only one element in row j is non-zero, and is

$$\max S_j^* = 1. \tag{18}$$

When all elements in row j are equal in absolute value—the worst possible situation—S_j^* is a minimum, and is

$$\min S_j^* = 1/q. \tag{19}$$

Let us redefine S_j^* as

$$S_j = a + bS_j^* \tag{20}$$

choosing a and b so that S_i varies from zero to one (rather than S_i^*'s varying from $1/q$ to 1). For this purpose we have the simultaneous equations in a and b

$$a + b\,(\max S_i^*) = a + b(1) = 1, \tag{21}$$

and

$$a + b\,(\min S_i^*) = a + b(1/q) = 0, \tag{22}$$

yielding

$$a = -1/(q - 1), \tag{23}$$

and

$$b = q/(q - 1), \tag{24}$$

from which

$$S_i = -\frac{1}{(q-1)} + \frac{q}{(q-1)} \frac{\sum_s v_{is}^4}{\sum_s (v_{is}^2)^2}. \tag{25}$$

Elementary algebra yields

$$S_i = (IFS(J))^2 = \frac{q \sum_s v_{is}^4 - (\sum_s v_{is}^2)^2}{(q-1)(\sum_s v_{is}^2)^2}. \tag{26}$$

as before. For the entire pattern matrix we can again derive the overall squared IFS as

$$(IFS)^2 = \frac{\sum_i [q \sum_s v_{is}^4 - (\sum_s v_{is}^2)^2]}{\sum_i [(q-1)(\sum_s v_{is}^2)^2]}. \tag{27}$$

Thus, all roads lead to Rome: regardless of what version of the quartimax criterion we use, if we set out to define an index which varies from zero to one we find (7), (14), or (26) for a given row j, or (10), (16), or (27) for the entire factor pattern matrix.

The Problem of Calibration

Although we know that an IFS can vary from a minimum of zero to a maximum of one, we still would like to know the meaning of values between these extremes. It would seem that only extensive numerical experience with real data will give us a solid feeling for, say, how big is big, etc. However, the following considerations may be helpful for this problem.

Above we uniformly defined *squared* indices of factorial simplicity. The reason for this is that throughout we have dealt with fourth-power functions of loadings, and it would seem more consistent with other statistical measures, *e.g.*, the correlation ratio, to deal with squared functions. Additionally, numerical appraisal indicates that, were we not to take square roots above, our indices would be too dramatically depressed for only slight departures from unifactoriality.

Further help in the calibration of our index may be gained from considering the special case of a row of a factor matrix with q elements, c of which are non-zero and equal in absolute value, and $(q - c)$ are zero (c is what is usually called the complexity of a variable represented by the row). Elementary analysis for this special case yields

$$(IFS(J))^2 = \frac{q - c}{c(q - 1)} \tag{28}$$

and

$$\lim_{q\to\infty} (IFS(J))^2 = \lim_{q\to\infty} \frac{q - c}{c(q - 1)} = \frac{1}{c}. \tag{29}$$

In Table 1 are given values of $IFS(J)$, as given by the square roots of (28) and (29), for various values of q and c. This table might possibly give the wrong impression about what IFSs are observed with real data. Most real world problems have an average complexity c of, so to speak, about one and one-half, and thus yield an IFS in the .70s or .80s.

Subjective reflection, based upon Table 1 and primarily observing IFSs for a substantial number of factor analyses from the real world, suggests the following evaluation of levels of our index of factorial simplicity:

in the .90s, marvelous
in the .80s, meritorious
in the .70s, middling
in the .60s, mediocre
in the .50s, miserable
below .50, unacceptable

Table 1

Indices of Factorial Simplicity as a Function of q, the Number of Factors, and c, the Complexity of a Variable's Loadings

c = complexity	2	3	4	5	6	7	8	9	10	11	12	∞
5				.0000	.2000	.2582	.2928	.3162	.3333	.3464	.3568	.4472
4			.0000	.2500	.3162	.3536	.3780	.3953	.4082	.4183	.4264	.5000
3		.0000	.3333	.4082	.4472	.4714	.4880	.5000	.5092	.5164	.5222	.5774
2	.0000	.5000	.5774	.6124	.6325	.6455	.6547	.6614	.6667	.6708	.6742	.7071
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

q = number of factors

Referring to Table 1, we see that an *IFS* will be at .50, the borderline of acceptability, when, for complexity two, there are three factors, for complexity three, there are nine factors, and for complexity four, there is an infinity of factors. For complexity c of five or greater, an acceptable *IFS* is not attainable.

REFERENCES

Carroll, J. B. An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 1953, 18, 23–38.

Saunders, D. R. An analytic method for rotation to orthogonal simple structure. *American Psychologist*, 1953, 8, 428 (Abstract).

Wrigley, C. & Neuhaus, J. O. The quartimax method: An analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology*, 1954, 7, 81–91.

Manuscript received 2/5/73
Revised manuscript received 9/5/73