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

Comments by Larry Hubert & Willem Heiser

6043 citations as of 4/1/2016

- **New Use for Quartimax**
- **Calibration of the Scale of the Index**
- **Remarkable Citations**

Factorial Simplicity Defined & Return to Quartimax

"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 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." (o.c., p. 31)

Comment: he is going to propose an index which measures tendency towards unifactoriality for **each row of the factor matrix** separately and for the matrix as a whole. That is exactly what the **quartimax criteria** intend to do.

Note that he rejected quartimax for rotation, but now finds a new use for it in measuring unifactoriality.

Three Criteria are Shown to be Equivalent if Rescaled

He takes the different quartimax criteria for analytic rotation of Carroll (1953), Wrigley & Neuhaus (1954), and Saunders (1953) and demonstrates mathematically that they become equivalent if we **scale them to lie between zero and one**.

He remarks:

> "It should be pointed out that, while we develop our index from the quartimax viewpoint, our results are applicable to any factor pattern matrix." (o.c., p. 31)

Hence the resulting index can be used:

- to compare quality of different rotations of one basic FA;
- to characterize quality of solutions for different datasets.

The Index of Factorial Simplicity (*IFS*)

Specifically, the resulting index *IFS* is defined as

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

$$(27)\qquad (IFS)^2 = \frac{\sum_j \left[q \sum_s v_{js}^4 - \left(\sum_s v_{js}^2\right)^2\right]}{\sum_j \left[(q-1)\left(\sum_s v_{js}^2\right)^2\right]}.$$

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.

Worst possible value of *IFS* is zero: occurs when **all loadings are equal** in absolute value for all rows; best possible value is one: occurs when **only one loading in each row is non-zero**.

The Problem of Calibration

Kaiser realizes that it would be useful to **calibrate** the *IFS* scale. To approach a reasonable answer, he considers 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. Here c refers to the **complexity** of the variable. He then obtains the following result:

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

So for a large number of factors, *IFS* converges to 0.50 for complexity 4, which he considers **too far** from unifactoriality.

Verbal Qualifications of the IFS Scale

By also considering *IFS*s for a substantial number of factor analyses "from the real world", he reached the following —**admittedly still subjective**—evaluations of levels in the *IFS* scale:

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

Unfortunately, almost all but a very small handful of the citations to Kaiser (1974) appear to be in error!

Some possible reasons for these misattributions are laid out in what follows.

The Kaiser-Meyer-Olkin Measure of SamplingAdequacy: The Root of the Problem

In his Psychometric Society Presidential Address in 1970 ("A second generation Little Jiffy," *Psychometrika*, 1970, *35* 401–415), Kaiser added a few embellishments to his well-known approach to factor analysis that Chet Harris had disdainfully labeled "Little Jiffy":

"principal components with associated eigenvalues greater than one followed by normal varimax rotation."

One such addition in this 1970 paper was a "measure of sampling adequacy" (MSA) that was intended to reflect whether it was reasonable to proceed with a factor analysis in the first place.

Kaiser attributed this MSA to work he was doing at the time with Professors Meyer at Loyola (Chicago) and Olkin at Stanford.

This MSA is now commonly referred to as the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO-MSA).

It is calculated routinely, for example, in the heavily-used SPSS and SAS factor analysis programs; also, functions for KMO-MSA appear in R, such as in the **psych** package developed by Bill Revelle.

The particular KMO-MSA computed by SPSS and various R functions, is not the exact same measure given in Kaiser (1970); instead, it is a modification meant to improve stability given in a 1974 Kaiser and Rice paper in *Educational and Psychological Measurement* ("Little Jiffy, Mark IV," *34*, 111-117).

The Citation Mixup Arises

Over the last four or so decades, when the value of the KMO-MSA index is reported (usually taken directly from SPSS output), the reference given for it is most often Kaiser (1974), the factorial simplicity paper, and not the correct citation of Kaiser and Rice (1974), the Version IV Little Jiffy article.

One reason for Kaiser (1974) becoming the inappropriately highly-cited paper it has developed into, may be due in part to an article by Charles Dziuban and Edwin Shirkey from *Psychological Bulletin*, also in 1974: "When is a correlation matrix appropriate for factor analysis" (*81*, 358–361).

the following sentence appears on page 359: "Kaiser's (1974) present calibration of the [KMO-MSA] index is as follows:"

No reference, however, appears for Kaiser (1974) in the bibliography for Dziuban and Shirkey (1974) but one is given for an "in press" piece by Kaiser and Rice.

The mixup of using Kaiser (1974) for Kaiser and Rice (1974) may be a miscitation phenomenon that is difficult to correct.

Two popular SPSS-related user manuals, for example, make this citation error: the *SPSS Survival Manual* by Julie Pallant, and *Discovering Statistics Using SPSS* by Andy Field; the same citation error is also made in other books by Andy Field, such as in *Discovering Statistics Using SAS*.

In hindsight, it is surprising that such a miscitation wasn't caught earlier by an author of these secondary SPSS user manuals.

Not a single statistical package (SPSS, SAS, SYSTAT, Matlab, or R) computes Kaiser's index of factorial simplicity, irrespective of how good or bad the index might be.

Several statistical packages do, however, compute KMO-MSA (SPSS, SAS, and Bill Revelle's **psych** package for R).

Comments by Bill Revelle

When Bill Revelle was asked why he didn't include a function for Kaiser's index of factorial simplicity in his **psych** R package, he said that he instead included a function for what he thought was a better index (of factorial complexity) due to Richard Hofmann ("Complexity and simplicity as objective indices descriptive of factor solutions," *Multivariate Behavioral Research*, *13*, 1978, 247–250).

He also noted that he was never tempted to use Kaiser (1974) in reference to his KMO-MSA function; he gave the three (correct) citations for it of Kaiser (1970), Kaiser and Rice (1974), and Dziuhan and Shirkey (1974).