Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187-200. (5237 citations according to Google scholar)

Applications of Principal Components Analysis or Exploratory Factor Analysis in the behavioral sciences are usually followed by rotation aiming at simple structure. And the rotation method used is very often Kaiser's normal varimax. What makes varimax so popular?

Varimax rotation is certainly not the only simple structure rotation method, neither was it the first. "The first analytic criterion for determining psychologically interpretable factors was presented in 1953 by Carroll". Kaiser thereby ignored earlier proposals to *partially* use analytic rotation criteria along with graphical methods (e.g., Tucker, 1944), probably because he strongly disregarded graphical methods: "… before the advent of computers, rotations were always carried out on a subjective graphical basis. Scientifically, of course, this was nonsense, and perhaps led, more than anything else, to a bad name for factor analysis among professional mathematicians and statisticians." (Kaiser, 1960, pp.146-147). Kaiser considered Carroll's proposal a breakthrough: the first method to optimize a single criterion operationalizing Thurstone's (1947) rules for simple structure. On the other hand, Kaiser (1958) made it clear that there was room for improvement, to say the least: "In the light of later developments, Carroll´s criterion should probably be relegated to the limbo of 'near misses'; however, this does not detract from the fact that it was the first attempt to break away from an inflexible devotion to Thurstone's ambiguous, arbitrary, and mathematically unmanageable qualitative rules for his intuitively compelling notion of simple structure." Almost simultaneously, Saunders (1953), Neuhaus and Wrigley (1954), and Ferguson (1954) proposed differently formulated, but equivalent methods for orthogonal simple structure rotation. Kaiser (1958) elegantly and briefly showed them to be equivalent to Carroll's method. Following Neuhaus and Wrigley, the method was called quartimax.

Kaiser (1958) observed that quartimax has a bias. To explain what tends to go wrong, he showed that quartimax was equivalent to maximizing yet another criterion, that is, the sum of rowwise variances of squared loadings. He wrote "…its aim is to simplify the description of each row, or test, of the factor matrix. It is unconcerned with simplifying the columns, or factors, of the factor matrix (probably the most fundamental of all requirements for simple structure). The implication of this is that the quartimax criterion will often give a general factor. …. there is no reason why a large loading for each test may not occur on the same factor." (Kaiser, 1958, p.190). And from these sentences, almost automatically the suggestion to maximize the varimax criterion emerged: If simplicity of the *columns* is what is most important, then one should maximize the sum of *column*wise variances of the squared loadings, that is, the varimax criterion. The introduction of varimax in Kaiser's 1958 paper was less dramatic than it might seem: Kaiser simply identified the sum of columnwise variances of the squared loadings as the (raw) varimax criterion he had presented *earlier* at the 1955 APA conference, while referring to its abstract (Kaiser, 1955). Also in 1955 the varimax criterion had been proposed in his unpublished PhD dissertation (see Carroll, 1957, or Kaiser 1979). The 1958 paper, however, is the first journal publication on varimax.

Kaiser (1958) offered quite a bit more than just the varimax criterion. He started out by comparing varimax rotation to quartimax rotation on two examples. He demonstrated the superiority of varimax over quartimax in that varimax didn't yield such a strong general factor as quartimax did. Nevertheless, he observed in the second example that also after varimax rotation, the variance contributions of the four factors differed considerably (though less than after quartimax). Kaiser guided us in his search for what caused this difference. "It seems reasonable to attribute the systematic bias […] to the divergent weights which implicitly are attached to the tests by their communalities. When one deals with fourth-power functions of factor loadings, a test with communality 0.6, for example, would tend to influence the rotations four times as much as a test whose communality was 0.3." and "There seems no rational basis for choosing among different weighting schemes. Let us then make the agnostic confession of ignorance which pervades any form of correlational analysis. For the purpose of rotation, weight the tests equally, in the sense that the lengths of the common parts of the test vectors have equal length. (The author is indebted to Dr. D.R. Saunders for this suggestion)." This rescaling leads to the currently prevalent "normal varimax" criterion. Kaiser demonstrated on an example that this procedure indeed had the effect of reducing the differences between the variance contributions of the rotated factors.

After his *ad hoc* introduction of the weighting scheme, Kaiser came with an interesting additional rationale. He proved mathematically, for the case of two factors, that, when tests in a cluster of tests all are collinear as far as their common parts are concerned, then "…*the normal varimax solution is invariant under changes in the composition of the test battery*". Such invariance does not hold for (raw) varimax or quartimax. It looks like a very strong property, and even though he admitted that his result holds only in a special case, he goes on to write that "The principle of simple structure may probably be considered incidental to the more fundamental concept of factorial invariance." He next verified what happens for a real data example, when adding variables (tests) one by one, and observed a high degree of stability of loadings on factors that are sufficiently well represented. However, one example of course does not prove a general invariance property. In fact, it is easy to demonstrate numerically that his result for the two factor case, does not generalize to cases of more than two clusters of collinear variables (and hence cases of more than two factors). Moreover, it can be shown that his proven result is simply a consequence of the fact that, in his special case, variances are considered of vectors with only two different values, and, such variances do not depend on how often such values occur. Hence, his invariance property holds for optimizing any criterion using columnwise variances, even for the opposite method of *minimizing* rather than maximizing the varimax criterion. So it seems that the factorial invariance property of normal varimax, even in this special case, is of no special value.

The paper essentially closes by a brief section on the oblique case. In these 14 lines, Kaiser proposed an alternative criterion for oblique rotation (which was dubbed covarimin by Carroll (1957) even before Kaiser published his paper), and he suggested how to optimize it. It has by no means become as popular as varimax, but it is remarkable that in this paper on orthogonal simple structure rotation, almost in passing, he offered a new criterion for oblique rotation as well.

The actual last pages of the paper concern a technical appendix, which describes a planar rotation procedure for iteratively maximizing the varimax criterion. Surprisingly, the paper does not finish with a discussion section, which might have given suggestions for future research or further developments. In hindsight, we can conclude that such a suggestion would have been superfluous. The usefulness of varimax apparently was self-evident, and many future developments have taken Kaiser's varimax as a starting point. One of the earliest such developments was Harris and Kaiser's (1964) successful "orthoblique" approach, where orthogonal varimax rotation was, in an ingenious way, used to find simple loadings for oblique factors. As another early follow up, the varimax solution was proposed to be used for setting up a simple target for oblique target rotation. The ensuing Promax method (Hendrickson & White, 1964) is also one of the more successful approaches to oblique simple structure rotation. Since then, various further developments along these lines have been proposed.

The varimax criterion has also inspired some authors to set up families of simple structure criteria, of which quartimax and varimax were the most prominent members (e.g., Clarkson & Jennrich, 1988). Furthermore, special procedures for application of varimax in other situations than that of a single ordinary factor loading matrix have been proposed. For instance, for the situation with several loading matrices for the same set of variables, Bloxom

(1968) proposed a method for rotation of these loading matrices to maximum similarity and maximum simple structure in the varimax sense.

To conclude, with varimax, Kaiser has offered a tool which has been and still is very successful. Its success not only becomes apparent in the practice of exploratory factor analysis, but also in the development of new rotation methods, in which varimax was used as a basis.


Bloxom, B. (1968) Factorial rotation to simple structure and maximum similarity, *Psychometrika, 33*, 237-247.

Carroll, J.B. (1953) An analytic solution for approximating simple structure in factor analysis. *Psychometrika, 18*, 23-38.

Carroll, J.B. (1957) Biquartimin criterion for rotation to oblique simple structure in factor analysis. *Science, 126*, 1114-1115.

Clarkson, D.B. & Jennrich, R.I. (1988) Quartic rotation criteria and algorithms. *Psychometrika, 53*, 251-259.

Ferguson, G.A. (1954) The concept of parsimony in factor analysis. *Psychometrika, 19*, 281-290.

Harris, C.W., & Kaiser, H.F. (1964) Oblique factor analytic solutions by orthogonal transformations. *Psychometrika, 29*, 347-362.

Hendrickson, A.E., & White, P.O. (1964) PROMAX: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology, 17*, 65-70.

Kaiser, H.F. (1955) An analytic rotational criterion for factor analysis. *American Psychologist, 10*, 438 (abstract).

Kaiser, H.F. (1960) The application of electronic computer to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Kaiser, H.F. (1979) This Week's Citation Classic, *Internal Bulletin University of Illinois, Urbana, IL, 24*, 11, 1979, 33.

Neuhaus, J.O. & Wrigley, C. (1954) The quartimax method: An analytic approach to orthogonal simple structure. *British Journal of Mathematical and Statistical Psychology, 7*, 81-91.

Saunders, D.R. (1953) *An analytic method for rotation to orthogonal simple structure.* Research Bulletin, RB 53-10. Princeton, New Jersey: Educational Testing Service.

Thurstone, L.L. (1947) Multiple factor analysis. Chicago: University of Chicago Press.

Tucker, L. R. (1944) A semi-analytical method of factorial rotation to simple structure. *Psychometrika, 1944*, 43-68.