PSYCHOMETRIKA—VOL. 35, NO. 4
DECEMBER, 1970

A SECOND GENERATION LITTLE JIFFY*

HENRY F. KAISER

UNIVERSITY OF CALIFORNIA, BERKELEY

For the past nine years I have had the very good fortune to be a member of a "Working Group in Factor Analysis." This group, under the leadership of Ledyard R Tucker, and sponsored by the office of Naval Research, has met semi-annually to talk, present papers (usually in preliminary form), and fight about matters factor-analytic. While I would not go so far as to say that everything done in these meetings has turned to pure gold, one has only to look at the journals during the sixties to see that Tuck's Working Group has been productive indeed. For me, at the very least, these sessions have provided a wonderful opportunity to come out of my hole, hear new ideas, have precious thoughts blasted, and go away with my batteries massively recharged.

Thus, let me dedicate my remarks to this group: Bargmann, Harman, Harris, Horst, Humphreys, Joreskog, Meredith, Tucker, Wrigley, together with occasional visitors, *e.g.*, Guttman, Olkin. If I am fortunate enough to say anything of value today, it surely stems from my association with Tucker and his troops.

In our spring meeting, 1964, at Paul Horst's shop in Seattle, we were in a particularly esoteric mood, or, I should say, the topics under discussion were particularly esoteric. This prompted Horst to ask about factor analysis on the firing-line, *i.e.*, boys, what do we usually do in the real world (as opposed to doing the fancy stuff under discussion)? My response to this question was to allude to that Sweet Young Thing, with downcast head and sorrowful eyes, who often enters my office with a large score matrix, neatly written out, and chirps despondently, "Just what *do* I *do*?" Reaching into the second right-hand drawer of my desk I pull out, usually with no further ado, that old nostrum, principal components with associated eigenvalues greater than one followed by normal varimax rotation, a procedure first proposed in the doctoral dissertation of an undistinguished education student 'way back in 1956 [Kaiser, 1956]. The Sweet Young Thing would take the prescription away, consult a computer, and several days later appear with cheeks aglow, happy as

---

* Presidential address delivered at the annual metting of the Psychometric Society Miami Beach, September 7, 1970. This research was supported in part by the Committee on Basic Research in Education (Patrick Suppes, Chairman), United States Office of Education, and the Committee on Research (Robert R. Brown, Chairman), University of California, Berkeley.

a lark, quite convinced that all sorts of dandy new things had come to pass. Nothing profound, perhaps, but at least some preliminary order out of well-perceived chaos.

Chester Harris' reaction to this proposed procedure was stark and simple: "Little Jiffy." To this day I am not sure if Chet was being disdainful of such an elementary method (in the presence of the abstruse stuff we were considering); I suspect, though, that he wasn't. In any case Tuck's Working Group no longer had to refer to "principal components with associated eigenvalues greater than one followed by normal varimax rotation;" we simply had to say "Little Jiffy." It is perhaps worthwhile for all of us to have this simple name, for Little Jiffy appears undoubtedly to be the most popular method of factor analysis extant, even to this day [Cronbach, 1970; Glass & Taylor, 1966].

It is traditional, in this talk, for the speaker to attempt to exhibit a little wisdom—to do a little big-picture philosophizing. Those of you who know me can testify that I would be the last person to defy tradition. So let me try briefly to give justification, for many problems, for the routinized, almost unthinking application of a factor-analytic method like Little Jiffy.

John Tukey, in talking about data analysis, has made the useful distinction between exploratory and confirmatory study. In confirmatory work, we know, a priori, or we think we know, a great deal about what will go on, and we are prepared to state rather sharp hypotheses about how the data will act. And our purpose is to adjudicate, in depth, these rather specific ideas. In exploratory work, on the other hand, we know very little ahead of time; the best we can do is take observations on a whole pot-full of random variables which we suspect may be relevant, and then see what happens. We're engaged, as it were, in a fishing expedition. I would suggest that factor analysis, as it is typically used in practice, is much more often exploratory than confirmatory; indeed, it often is not even exploration-in-depth, it is reconaissance, to use another of John Tukey's distinctions.

We can envision a continuum of kinds of studies, running from pure exploratory on the left to pure confirmatory on the right:

exploratory |————————————| confirmatory

I have found it fun, and perhaps even useful, to scale the interests of methodological factor analysts along this continuum. (Any of you who want your scale value, check with me after the meeting.) Engaging in a little self-analysis, I find that, on this continuum, I am an extreme left-winger. And, among methodologists, if not practitioners, I'm a little lonely there. For most of my thoughtful colleagues, when they speak of fishing expeditions, are speaking disdainfully. To use L. J. Savage's [1954] distinction, they prefer to "look before they leap" rather than "cross that bridge when they come to it." My thought is that it is often not possible to look before you leap—sometimes, if you want to cross a stream, it is necessary to leap blindly, often landing in the stream and doing a good deal of thrashing about before getting to the other side.

Consider, for example, a study of the psychasthenic aspects of socialistic infiltration into the administration of Rutherford B. Hayes. Little is known; the best we can do is make a yes-no decision about which variables are relevant. Following Clemenceau, age is relevant. Most folk would say that length of right forearm is not. Gathering together the relevant variables, or a (psychometric) sample of them, we take many observations, subject them to something like Little Jiffy, and almost certainly come away wiser, if not wise.

I should point out, of course, that whether to explore or whether to confirm is not competitive, that a given study "naturally" falls somewhere along the continuum given above. I am only suggesting that many, many problems can be of necessity only exploratory, and it behooves us to look at the methodological side of things for such problems.

I said above that I was lonely among my immediate pals, the methodologists, but I am *not* lonely among the practitioners, for, again, it would seem that the majority of factor-analytic studies are exploratory, or nearly so. Again, I guess, I qualify as a left-winger, for I want to *do good* for the general (factor-analytic) public.

And that is my purpose today. Little Jiffy, that seemingly reliable workhorse for exploratory studies, is old, and in need of an overhaul. There has been a great deal of important work done in the last fifteen years which, if pasted together, hopefully will produce "A Second Generation Little Jiffy." Interestingly, the sort of gluing that needs to be done is almost exactly the same as it was for the original Little Jiffy. For that affair, I had to answer the number-of-factors question for Hotelling's [1933] principal components, and then I had to apply a mild polish to Charles Wrigley's [Neuhaus & Wrigley, 1954] ground-breaking quartimax transformation method. Today, my principal proposals will again be to suggest an answer to the number-of-factors question for factors developed according to Harris' [1962] monumental 1962 *Psychometrika* paper, and then, at the transformation stage, apply a mild polish to Harris' elegant orthoblique method [Harris & Kaiser, 1964].

As a final preliminary, let me announce two principles which I found useful in the present work. The first of these is:

Principle I. It don't make no never-mind.

The thought here is that when faced with a crucial decision, don't try to settle it; rather, avoid it! An example which comes to mind, in factor analysis, is the question of the "real" or "true" scale for the variables under consideration. Rather than fuss and fight about it, avoid the question by using one of the several *scale-free* methods which have been proposed, *e.g.*, canonical factor analysis [Rao, 1955], alpha factor analysis [Kaiser & Caffrey, 1965], etc.

The second principle is:

Principle II. Simplicity is elegance; elegance is simplicity.

The thought here is, if possible, always to use the simplest, most straightforward mathematics, without special qualifications. For example, the famous Swiss mathematician, Leonhard Euler, at the Court of Catherine the Great, announced, with pristine eloquence

$$x = \frac{a + b^n}{n},$$

thus proving that God exists. Dealing with a problem of nearly as great importance, the German physicist Einstein looked sternly into the middle distance, and said, simply

$$E = mc^2.$$

Examples from factor analysis would be that we should prefer principal axes before centroids, quartimax before varimax.

Measuring Sampling Adequacy of Factor-Analytic Data Matrices

Let us use the term *factor analysis* generically as an expression for what Maurice Kendall [1950] calls the analysis of interdependence. For exploratory factor analysis, then, three formal mathematical models have been proposed. They are: (1) component analysis (Hotelling [1933]), (2) common factor analysis (Thurstone [1947]), and (3) image analysis (Guttman [1953]). (Thus, I have used, and shall use, the term *factor* generically to mean a component, a common factor, or an image factor.)

For our problem I have indicated elsewhere [Kaiser, 1958a, 1963] that, of these three, I prefer Guttman's image analysis, primarily because it seems an ideal compromise between component analysis and common factor analysis: according to Principle II, like component analysis, it is mathematically simple and straightforward (yielding, for example, factor scores exactly), while approaching, psychometrically (*i.e.*, as the number of variables p goes to infinity) the strict, but incredibly complex, common factor analysis of Thurstone. Among Guttman's many landmark papers [Guttman, 1940, 1953, 1954, 1956] on image analysis and related topics one particular theme appears repeatedly: given a correlation matrix R we should always look carefully at R^{-1}, in order to assess the sampling adequacy of the data for factor-analytic purposes. Guttman has proved that the matrix R^{-1} should be near-diagonal for factor analysis to be an appropriate tool. For more than fifteen years now I have observed, with real data, the practical truth of

Guttman's theoretical assertions. And for at least a dozen years I have tried to explicate Guttman's general finding into a simple formula involving R^{-1}, in an attempt to measure the sampling adequacy of factor-analytic data matrices. Four distinct earlier efforts ultimately failed—they just did not always yield what obviously was the "right" answer for some problems. During the last two years, working with Professor Meyer of Loyola, Chicago, and Professor Olkin of Stanford, we appear to have succeeded in getting something which has yet to fail in any obvious way. Consider $Q = SR^{-1}S$, where $S = (\operatorname{diag} R^{-1})^{-1/2}$. Q is the anti-image intercorrelation matrix. Then consider the function

$$MSA = 1 - \frac{\sum\sum_{j \neq k} q_{jk}^2}{\sum\sum_{j \neq k} r_{jk}^2}$$

as a measure of sampling adequacy, a relative measure of the "amount" of correlation, before and after Guttmanization. A similar measure

$$MSA(J) = 1 - \frac{\sum_{\substack{k \\ k \neq j}} q_{jk}^2}{\sum_{\substack{k \\ k \neq j}} r_{jk}^2}$$

may be defined for each variable separately. $MSA(J)$ measures to what extent a given variable "belongs to the family," psychometrically. Now any MSA clearly lies between minus infinity and plus one. In practice, for just any old correlation matrix from the real world, the over-all MSA apparently will always be positive, and almost invariably be above .40 or .50. It appears that we don't have good factor-analytic data until MSA gets to be at least in the .80s, and really excellent data does not occur until we reach the .90s.

MSA appears to be a function of four "main effects." Holding the others constant, MSA improves as the number of variables p increases; MSA improves as the (effective) number of factors q decreases; MSA improves as the number of subjects n increases; and MSA improves as the general level of correlation $\bar{r}$ increases. I suggest that these main effects all go in the right direction.

These results were obtained from a number of random applications plus a systemative application to the 54 correlation matrices of an extensive recent study of Tucker, Koopman, and Linn [1969]. After obtaining these results I asked Professor Tucker for a subjective rating of the quality of the data in his 54 correlation matrices (not telling him, of course, the results we had obtained). The rank correlation between Tucker's subjective ratings and the Kaiser–Meyer–Olkin MSA described above was .85. May I say that I think we're on to something here—that to correlate .85 with the judgment of the World's Wisest Psychometrician is worth writing home about?

Allow me to say that our fussing to find an appropriate measure of sampling adequacy for factor-analytic data matrices is a good illustration of the distinction between exploratory and confirmatory study. At the outset, confronted only with the stark Guttmanian knowledge that R^{-1} should be near-diagonal, I fussed and fumed in the most idiotic (in retrospect) fashion. Now that we have begun really to zero-in on the problem, we have only to confirm, with "well-designed" study, the details of our ideas. I cannot resist saying that, for me at least, the earlier exploratory thrashing about was much more fun—and perhaps even represented more progress—than the forth-coming confirmatory prettying-up.

Initial Factoring

Let us say that we have data with a respectable MSA. How, then, do we factor it? If I am to be consistent with my earlier expressed preference for image analysis, I should find, say, the principal axes of Guttman's image covariance matrix. But in maintaining my preference for image analysis, I am turning my back on Hotelling and Thurstone, and I find that most unpalatable. Wouldn't it be dandy if, in some fashion, it were possible to factor in a way which was independent of the major distinctions among these three models? For we would be applying Principle I—it don't make no never-mind.

Well, it turns out that we can apply Principle I. Just reach for *Psychometrika*, 1962, and read the most important factor-analytic paper of the last decade: Chester W. Harris [1962], "Some Rao–Guttman Relationships." For Harris has shown that by rescaling in the metric of the anti-images—with S^{-1}—the R of component analysis, the $R - S^2$ of common factor analysis, and the $R + S^2R^{-1}S^2 - 2S^2$ of image analysis, we not only get solutions which are scale-free, we get matrices all of which have the *same* eigenvectors E (although different eigenvalues; the eigenvalues are M^2, say, $M^2 - I$, and $M^{-1}(M^2 - I)^2M^{-1}$ for the three models, respectively). The important implication of this (although it's not trivial to prove) is that having the same eigenvectors implies having the same factors. (I should say that the factoring of $R - S^2$ with Harris rescaling is only the "natural" first approximation to the strict maximum-likelihood Thurstonian common factor solution of Lawley [1940] and Rao [1953]). Thus, we may say that Harris factors are Harris factors, regardless of the model; they are, if I may introduce a new term, *model-free*. We have applied Principle I with a vengeance. Thanks to Harris, it is clear that Hotelling, (the first approximation to) Thurstone, and Guttman are now buddy-buddy, model-wise.

In passing, let me note another goodie which may obtain from Harris factoring. Let us say that some or all of our variables are dichotomous or binary, *e.g.*, test items scored zero and one. It is well-known in the real world that such data may lead to artifactual "difficulty factors." What to do about this, at least theoretically, has proved a messy problem. In the present context we are considering images, the linear least-squares predicted value of one variable from the remaining $p - 1$. Even if all the variables are dichotomous, a crude appeal to the Central Limit Theorem suggests that the images will be sensibly multivariate normal, a set-up which is well-known *not* to produce difficulty factors. It may well be, then, that what I am proposing today circumvents this long-standing problem. Further work is indicated.

On the Number of Harris Factors

Let us turn now to what is the toughest question we have to answer. And that is the number-of-factors question: how many Harris factors should be retained?

Over the years I have noticed that one can retain too many factors, without doing any real harm, under special conditions. These conditions for robustness are that we use either image analysis or common factor analysis and that we ultimately obtain transformed ("rotated") factors which are uncorrelated—the so-called orthogonal case. What happens is that, under transformation, small garbage factors do not affect, in any substantial way, the transformed major factors. This observation fits in nicely with another one. The "natural" number of factors, in the context of Harris' paper, is Guttman's [1954] well-known stronger (algebraic) lower bound, the number of positive eigenvalues of $R - S^2$. This number, for real data sample correlation matrices, from the point of view of ultimate possible scientific meaning of the factors, is "too many"—invariably about three-fifths the number of originally observed variables [Kaiser & Hunka, 1960]. I have tried to put these two observations together, but have failed. The problem is that I do *not* want to maintain the restriction of uncorrelated factors; in particular I would like to use Harris' orthoblique at the transformation stage, and this method effectively gives each factor equal weight, which implies that the sort of robustness described above does *not* occur.

My next attempts to answer the question of the number of Harris factors involved what Bert Green [1966] calls root-staring, plotting the value of eigenvalues against their ordinal numbers, consulting my tummy for a nice answer upon staring at this plot, and then seeing if I could make up some Principle II, rational, simple rule of behavior which explicates objectively what my tummy tells me. Four million trials failed: anything that I was able to dream up would always give obviously bad answers for some problems.

Finally, I turned to the sort of thinking represented in the original Little Jiffy, where the rule of behavior for the number of factors is to retain those principal components of R with associated eigenvalues greater than one. (That same rule also applies, in common factor analysis, for my alpha factor analysis [Kaiser & Caffrey, 1965].) Now that rule of behavior can be rationalized in a number of ways [Kaiser, 1960]. What I have done is develop yet a new rationale for the number-of-factors in the original Little Jiffy, and then apply the same rationale to the present case of Harris factors.

Consider the vector representation of the original variables—the so-called "variable space." Each of the variables is given by a vector (all with a common origin), the length of a vector represents the standard deviation of the variable, and the cosine of the angle between two vectors gives the correlation between the two variables represented. Now place a unit-length vector, representing a factor, at random (but with the same origin) in the space spanned by the test vectors and consider the sum of the squares of the projections of the test vectors on the factor vector. And, finally, find the expected value of this sum-of-squares as the factor vector sweeps over the entire space. This can be a messy problem in iterated integration, but, furtunately, it can be solved by strictly algebraic means. For the original Little Jiffy, where the test vectors are all of unit-length, it turns out that the expected sum-of-squares of projections is one—a result which is almost intuitively obvious. Now if we allude to "greater than chance" type arguments, we would say that we should retain a factor if and only if the sum-of-squares of projections of test vectors on it is greater than one, thus establishing yet another rationale for the rule for the original Little Jiffy.

Let's do the same thing for Harris factors. Under the component analysis model, the test vectors are of length s_j^{-2}, invariably greater than one. The result here is that we should retain only those Harris factors with associated eigenvalues m_q^2 greater than the mean s_j^{-2}, which turns out to be those factors greater than the mean eigenvalue of $S^{-1}RS^{-1}$. The same result is obtained under the common-factor analysis model, but not the image analysis model; thus the result is only two-thirds model-free.

I have tried this rule on a goodly number of examples. And it has never failed in any obvious way. In cross-checking this rule with the rule for the original Little Jiffy, I found that invariably the present rule always gave a number of factors less than or equal to that of the original rule. I tried to prove this, and couldn't, so I bombed all my algebraic pals around the world with it, and they couldn't prove it either. Finally, Professor Stanley Mulaik of the University of North Carolina produced a counter-example to the conjecture, using an artificial simplex-like correlation matrix.

But I have yet to obtain a counter-example with real data. Perhaps the conjecture could be specialized a little and become a theorem.

While, as I indicated above, this Kaiser–Mulaik rule of behavior for the number of Harris factors has never come close to giving an absurd result, it may be a little on the conservative side, for some problems. I find this distressing, if true. I'm afraid that we shall have to wait and see as more results with real world data become available.

The Transformation Problem

Harris, with a minor assist from me [Harris & Kaiser, 1964], in 1964 developed what has become known as the "orthoblique" method for the transformation problem, so-called because of the use of orthonormal transformations to obtain factor-analytic solutions involving correlated factors. A special case of this method, and the one which will concern us here is to apply an orthonormal transformation T to the *unit-length* eigenvectors E. In general, this leads to correlated factors, despite the transformation's being orthonormal. This orthoblique solution, as it stands, is a definitive solution to the cluster analysis problem (at least the editors of our favorite journal let us get away with saying "definitive"!). For the general factor-analytic problem this solution has the property, apparently, of transforming the axes, always, in the right direction—the best simple structure can always be seen, although it may be distorted, and for data deviating substantially from a cluster-analytic pattern, the distortion is sufficient to call for action. Our problem today is then how best to put the final polish on Harris' orthoblique, just as earlier I polished up Wrigley's quartimax with varimax for the original Little Jiffy.

I should say that Harris' orthoblique as described here applies Principles I and II, passionately. First, in working on the unit-length eigenvectors E, which are the same under all three mathematical models, we are again applying Principle I—it don't make no never-mind—the transformation is model-free. Second, the orthonormal transformation matrix T is generated by (raw) quartimax so that we are applying Principle II in choosing quartimax, the simplest of the orthomax criteria (quartimax [Neuhaus & Wrigley, 1954], varimax [Kaiser, 1958b], equamax (Saunders, 1962], etc.). But, then, again, we are applying Principle I because in obtaining a quartimax solution we are also obtaining a varimax or any other orthomax solution (the term which distinguishes among these criteria, here, is, and remains, constant under transformation [Harris & Kaiser, 1964], thus disappearing under differentiation, and consequently it don't make no never-mind).

Putting the needed polish on orthoblique has been a problem with which I have struggled for six years. Almost all of my efforts were guided by Principle II so that I attempted to state some additional pure and simple analytic criterion for the job. It certainly seems much more pleasing, in one fell swoop, to write down some clean-cut to-be-optimized function which is easily differentiable and leads to the "right" answer. But all of my many efforts of this sort failed. Because I was determined, without bias, to solve Thurstone's famous box problem, of the form (see Figure 1). Even the most esoteric pure analytic criteria just couldn't handle this problem. So with

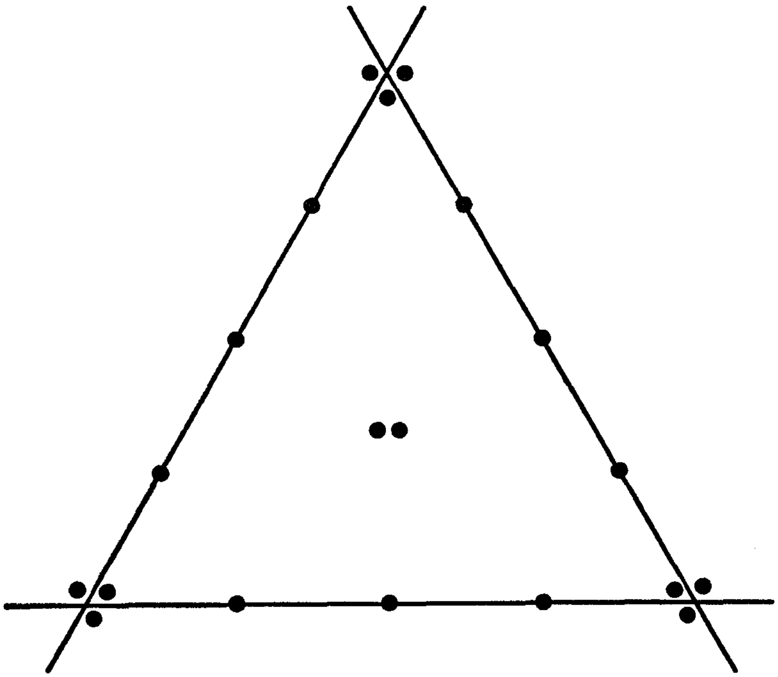

Figure 1
Thurstone's Box Problem

the deepest regret I abandoned Principle II and analytic criteria completely, and turned to the aesthetically less satisfying business of fitting hyperplanes. For this task, remember that I had the tremendous advantage of being 99% of the way via Harris' orthoblique.

I talked at some length with Professor Tukey about the problem and he remarked that it seemed similar to the problem of treating outliers in statistics. His thoughts and recommendations are detailed in his already classic "The Future of Data Analysis [Tukey, 1962]." For our purpose we need to distinguish between "trimming"—actually throwing away very bad test vectors—and "winsorizing"—changing or moving mildly bad test vectors. Since an orthoblique solution applies the varimax criterion, accordingly the "natural" cut-off point for trimming is to reject those test vectors with an orthoblique projection (in absolute value) greater than the root-mean-square projection for the factor under consideration. This rule, it appears, always gets rid of at least as many test vectors as it should, and usually leaves some which need attention. This attention could be further trimming, one at a time, say, but that provokes the seemingly unanswerable question of where to stop. It seems preferable to winsorize the remaining moderate outliers from—the hyperplane.

The way we have done this is as follows. After the initial trimming, fit a hyperplane to the remaining test vectors in a least-squares sense (it's an easy eigenproblem). Then look for the test vector whose terminus is furthest from the present hyperplane and plop it right down on that hyperplane, replacing it with its projection on the hyperplane. Then fit a new hyperplane, winsorize the most aberrant test vector into the new hyperplane, fit a new hyperplane, winsorize, etc., until convergence, *i.e.*, until all the winsorized and rewinsorized vectors lie precisely in the final hyperplane. Don't be confused: this winsorizing is only a temporary measure used to find optimally the transformation vector orthogonal to the final hyperplane, a transformation vector which is applied to the factor matrix with the test vectors back in their original places. Again, aesthetically, I am not too pleased with this procedure, but, by golly, it really works. In particular it handles the above box problem perfectly (obvious exact zero loadings come out exactly zero). Perhaps I would be happier if I were to expouse a new Principle saying, "It matters not how you get there; contentment comes from getting there, right on the button."

I might comment for those of you who may be puzzled by this that it is instructive to take a small sample of not-all-equal numbers and do the following. Find the one with the greatest deviation from the mean and change it to the mean; now find the observation furthest from the new mean and change it to this new mean, keeping it up until all the numbers are equal. Playing around with numbers in this fashion suggests that we are engaged in a complicated way of finding the mode—the point of greatest density. In the multivariate extension of this to the factor-analytic problem described above, we are also, it seems, finding the multivariate mode—the hyperplane of greatest test-vector density, which is what hyperplane fitting is all about.

One final comment about this Kaiser-Tukey winsorizing business: above when I said that we were 99% of the way with orthoblique, I was not using a figure of speech. In some 40 or 50 studies involving hundreds of factors the average correlation between an original Harris orthoblique factor and its winsorized counterpart was .99. It is clear that we have gone to a lot of trouble to apply a very mild polish.

The final step in the solution is to find the intercorrelation matrix of the factors. Since the eigenvalues of the original covariance matrix enter into the appropriate formula, and since these eigenvalues are not model-free, this matrix is not model-free. Thus, we have, at this final stage, had to abandon Principle I. However, some numerical work indicates that the differences in the intercorrelation matrices for the factors, according to which set of eigenvalues we use, are small, leading to the tentative assertion that our unfortunate last-gasp abandoning of Principle I is practically of little importance.

Similarly the factor-score matrices, according to the three models, are slightly different.

Computer Program and Example

Doctor John Wells of Berkeley's Computer Center has developed a program for our new Little Jiffy. This program is available by writing me at the School of Education, University of California, Berkeley.

The program outputs twenty things:

1. Original intercorrelation matrix.
2. Root-mean-square original correlation, for each variable on the remaining $p - 1$.
3. Overall root-mean-square original correlation.
4. Anti-image intercorrelation matrix.
5. Root-mean-square anti-image correlation, for each variable on the remaining $p - 1$.
6. Overall root-mean-square anti-image correlation.
7. *MSA*, for each variable on the remaining $p - 1$.
8. Overall *MSA*.
9. Squared multiple correlation of each variable on remaining $p - 1$; variance error of estimate of each variable on remaining $p - 1$.
10. Harris eigenvalues.
11. Mean Harris eigenvalue.
12. Number of factors.
13. Harris (column) eigenvectors.
14. Harris–Kaiser orthoblique transformation.
15. Orthoblique factor matrix.
16. Orthoblique common factor intercorrelation matrix,
17. Kaiser–Tukey winsorized orthoblique transformation.
18. Winsorized orthoblique factor matrix (scaled so that each column has mean square loading equal to one).
19. Winsorized orthoblique common factor matrix (conventionally scaled).
20. Winsorized orthoblique common factor intercorrelation matrix.

As an example, let us take Holzinger and Harman's classic 24 psychological tests. Skipping most of the above-mentioned output, we observe an overall *MSA* here of .86—good but not excellent (my educated guess is that the value is not larger because of the sample size of individuals being only $n = 145$; the psychometric sampling for the $p = 24$ variables is well-known to be excellent). The final common factor matrix, together with the intercorrelation matrix of the common factors is given in Table 1. Note that the number of factors here is four, in contrast to the original Little Jiffy's having five. For this famous problem Harry Harman (personal communication) answers unequivocally that the best number of factors for the problem is either four or five, but refuses to commit himself as to which of these two

TABLE 1

Holzinger and Harman's 24 Psychological Tests

	H & H Subjective Solution					Little Jiffy					
	A	B	C	D	h_j^2	A	B	C	D	SMC_j	MSA_j
1	70	-07	09	04	54	70	-03	06	05	51	89
2	46	-02	-01	02	21	49	-00	-02	00	30	81
3	60	02	-08	-01	32	63	00	-15	01	44	72
4	58	07	-00	-05	35	57	12	-05	-05	41	82
5	-04	83	09	-06	67	01	81	08	-07	67	86
6	-04	83	-08	09	66	-00	83	-09	05	68	88
7	-05	93	04	-12	74	03	89	01	-15	68	88
8	22	52	16	-08	52	26	55	10	-07	56	91
9	-08	90	-20	16	75	-02	88	-12	03	71	86
10	-32	08	78	08	57	-11	05	77	09	58	76
11	-18	07	61	24	53	00	02	51	34	54	82
12	14	-15	79	-12	54	26	-17	71	-00	54	80
13	35	01	64	-20	56	46	01	49	-07	54	88
14	-24	18	01	58	32	-11	11	00	57	36	82
15	-07	03	-10	61	29	-01	01	-01	55	29	86
16	31	-08	-07	47	38	35	-13	-07	52	43	88
17	-24	05	01	76	46	-07	-01	14	62	41	82
18	14	-19	13	59	46	25	-20	26	47	44	87
19	08	02	03	42	25	17	-00	07	38	37	80
20	37	25	-05	17	42	34	28	-05	19	46	93
21	32	-01	32	13	41	40	-01	35	12	47	90
22	26	28	-15	34	43	32	27	-03	21	45	92
23	49	21	08	08	53	49	24	08	09	56	90
24	-01	27	39	18	47	07	25	40	20	53	92
A	100	58	46	58		100	52	26	45		
B	58	100	46	51		52	100	36	52		
C	46	46	100	60		26	36	100	30		
D	58	51	60	100		45	52	30	100		

answers is better. The same matrices, as determined by Holzinger and Harman subjectively, are also given in Table 1; clearly, the correspondence between the two common factor matrices is excellent, between the common factor intercorrelation matrices not as good.

Concluding Remarks

I wish I could say that I've wrapped up the problems of exploratory factor analysis (something that my secret heart would like to do) but that would not only be immodest but untrue. I can only say that I've done the best I can and I hope, at the least, to have glued together the excellent work of Harris et al in a way which provides a basis for action in the world of real data. Let me conclude by reviewing the possible faults of the new Little Jiffy and by outlining some possibilities for future work.

My principal reservation about what I have developed today is the rule of behavior for the number of Harris factors. I'm fairly sure that occasionally (how occasionally, I don't know) this rule will lead to too few factors. And I suspect that errors of this sort may not be easy to detect because they will be small errors: Barry Goldwater conservatism, not the "conversatism" of George Wallace. I have exhausted myself with this problem, and thus regrettably have no very substantial alternative to suggest. Perhaps the provocative work of John Horn [1965], tailored to the present context, might be a possibility.

I would still like to see work on pure analytic transformation methods, transformation procedures more in keeping with Principle II than the winsorizing described above. For example, just very recently I have worked out some seemingly elegant mathematical properties of the old Kaiser–Dickman [1959] binormamin criterion; I only wish that the method worked better on really tough problems, *e.g.*, the box problem. But improvements might very well be possible.

But the most important future work, as I see it, should continue to concentrate on the number-of-factors question. I still think that the best answer to this is deliberate overfactoring—if only oblique transformational methods were robust under overfactoring. But, as far as we now know, they ain't.

For exploratory factor analysis it might be interesting to consider further the backwards sort of rescaling represented in alpha factor analysis and the normal orthomax solutions. This has been a major thrust in my own theoretical work [Kaiser, 1958b; Kaiser & Caffrey, 1965; Kaiser & Dickman, 1959] and I am not yet completely convinced that it is to be discarded for the statistician's rescaling in the metric of the unique parts, or for Harris' rescaling, in the metric of the anit-images, as I have done today.

For the future, I can assure you of one thing: if any of you folk, more talented than I, come up with some Big Break-throughs I shall be waiting in the wings ready and eager to paste them together to produce the next generation Little Jiffy.

REFERENCES

Cronbach, L. J. *Essentials of psychological testing.* (3rd ed.) New York: Harper and Row, 1970.

Glass, G. V and Taylor, P. A. Factor analytic methodology. *Review of Educational Research,* 1966, 36, 566–587.

Green, B. F. The computer revolution in psychometrics. *Psychometrika,* 1966, 31, 437–445.

Guttman, L. Multiple rectilinear prediction and the resolution into components. *Psychometrika,* 1940, 5, 75–99.

Guttman, L. Image theory for the structure of quantitative variates. *Psychometrika,* 1953, **18**, 277–296.

Guttman, L. Some necessary conditions for common-factor analysis. *Psychometrika,* 1954, **19**, 149–162.

Guttman, L. "Best possible" systematic estimates of communalities. *Psychometrika,* 1956, **21**, 273–286.

Harris, C. W. Some Rao-Guttman relationships. *Psychometrika,* 1962, **27**, 247–263.

Harris, C. W., and Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. *Psychometrika,* 1964, **29**, 347–362.

Horn, J. L. A rationale and test for the number of factors in factor analysis. *Psychometrika,* 1965, **30**, 179–185.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology,* 1933, **24**, 417–441, 498–520.

Kaiser, H. F. *The varimax method of factor analysis.* Unpublished doctoral dissertation, University of California, Berkeley, 1956.

Kaiser, H. F. *The best approximation of a common-factor space.* University of California Res. Rep. 25, Contr. No. AF 41(657)–76, 1958a.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* 1958b, **23**, 187–200.

Kaiser, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement,* 1960, **20**, 141–151.

Kaiser, H. F. Image analysis. In Harris, C. W. (Ed.) *Problems in measuring change,* Madison: University of Wisconsin Press, 1963, 156–166.

Kaiser, H. F., and Caffrey, J. Alpha factor analysis. *Psychometrika,* 1965, **30**, 1–14.

Kaiser, H. F., and Dickman, K. W. Analytic determination of common factors. *American Psychologist,* 1959, **14**, 425. (Abstract)

Kaiser, H. F., and Hunka, S. Some empirical results with Guttman's stronger lower bound for the number of common factors. Unpublished manuscript, 1960.

Kendall, M. G., and Babington-Smith, B. Factor analysis. *Journal of the Royal Statistical Society B,* 1950, **12**, 60–94.

Lawley, D. N. The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh,* 1940, **60**, 64–82.

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

Rao, C. R. Estimation and tests of significance in factor analysis. *Psychometrika,* 1955, **20**, 93–111.

Saunders, D. R. Trans-varimax: some properties of the ratiomax and equamax criteria for blind orthogonal rotation. *American Psychologist,* 1962, **17**, 395. (Abstract)

Savage, L. J. *The foundation of statistics.* New York: Wiley, 1954.

Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. Chicago Press, 1947.

Tucker, L. R, Koopman, R. F., and Linn, R. L. Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika,* 1969, **34**, 421–459.

Tukey, J. W. The future of data analysis. *Annals of Mathematical Statistics,* 1962, **33**, 1–67.