5

A History of Factor Indeterminacy

James H. Steiger
Peter H. Schönemann

Louis Guttman, a noted contributor to the psychometric literature in a number of areas, including factor analysis and reliability theory, succinctly summarized the basic nature of the indeterminacy problem in the following analogy (Guttman, 1972):

> The phenomenon is not peculiar to factor theory (of which $T + E$ theory is a special case). One of the difficulties in discussing it in the factor analysis

Note: We gratefully acknowledge the support by NIMH (Small Grant 1R03MH25680, NRC 9Grant No. A4640) and the University of British Columbia (UBC Humanities and Social Sciences Grant 26-9508).

context is that of disentangling the problem from extraneous details of the mathematics of factor analysis.

It may help to make the point clear by considering the following "true score" problem. Suppose it is known that for a certain variable T, individual i has one and only one value T_i, and this value satisfies

$$T_i^2 - 100T_i + 99 = 0 \tag{1}$$

What is individual i's score on T?

The answer, of course, is that T_i is either 99 or 1; the statement of the problem leaves the actual value indeterminate in the sense that too many solutions exist to condition (1). This is quite different from an *estimation* problem where *no* exact solution exists, and therefore one seeks an approximation. To say one should "estimate" T_i by taking an average of the two solutions (say the arithmetic mean 50) is to introduce an arbitrary step not inherent in the initial problem. Such an average may be good for some purposes and bad for others. If a further purpose were clearly defined, it might help decide between the actual solutions 1 and 99, and obviate any thought of averaging.

Such, I believe, is the gist of the phenomenon of indeterminacy. To understand it requires no expertise in factor theory.

Guttman's analogy clarifies some of the more confusing facets of factor indeterminacy, and we shall refer to his example frequently.

INTRODUCTION TO THE PROBLEM

For simplicity, we shall focus our introduction to factor indeterminacy on the sample case of the single factor model. However, the factor determinacy problem generalizes directly to multiple factor and population cases.

The Single Factor Model and Classical Test Theory. In 1904, Charles Spearman advanced the theory that the scores of a group of people on a number of tests can be explained in terms of a single underlying factor, called "general intelligence," which the tests measure in common. His single factor model, the mathematical statement of this theory, explains the test scores y_{ji} of each person i as a weighted combination of a single underlying factor x, common to all the tests, and a factor z_j specific to the test j. Thus, if y_{ji} is the score of person i on test j, the single factor model states:

$$y_{ji} = a_j x_i + u_j z_{ji} \tag{2}$$

with the restrictions

$$\sum_i x_i z_{ji} = 0 \tag{3}$$

for all j, and

$$\sum_i z_{ji} z_{ki} = 0 \tag{4}$$

for $j \neq k$.

The a_j's and the x_i's each convey significant information. The a_j's, which indicate the extent to which test j correlates with the underlying factor x, are frequently called "factor loadings." The x_i's, on the other hand, which show how much of the factor x is possessed by each person i, are generally referred to as "factor scores." The two restrictions in Equations (3) and (4) stipulate that the common factor is uncorrelated with all specific factors and that each specific factor is uncorrelated with every other specific factor.

In matrix notation, the single factor model can be written as

$$_n\mathbf{Y}_N = {}_n\mathbf{a}_1 \mathbf{x}'_N + {}_n\mathbf{U}_n \mathbf{Z}_N \tag{5}$$

where

$$\frac{\mathbf{ZZ}'}{N} = \mathbf{I}_n \tag{6}$$

and

$$\frac{\mathbf{X}'\mathbf{Z}'}{N} = \mathbf{0} \tag{7}$$

It is important to realize that the classical test theory model is a special case of this single factor theory. If ${}_n\mathbf{Y}_N$ contains the scores of N people on n parallel forms, the classical test theory model can be written:

$${}_n\mathbf{Y}_N = {}_n\mathbf{T}_N + {}_n\mathbf{E}_N \tag{8}$$

where

$$\frac{\mathbf{TE}'}{N} = \mathbf{0} \tag{9}$$

and

$$\frac{\mathbf{EE}'}{N} = (\sigma_e^2)\mathbf{I} \tag{10}$$

The test theory model partitions a person's test score into a true score component (t) and an error component (e), which is uncorrelated with the true score. The single factor model translates into the classical test theory model by setting

$${}_n\mathbf{T}_N = {}_n\mathbf{a}_1\mathbf{x}'_N \tag{11}$$

and

$${}_n\mathbf{E}_N = {}_n\mathbf{U}_n\mathbf{Z}_N \tag{12}$$

where

$$\mathbf{a}' = (a, a, a, \ldots, a) \tag{13}$$

and

$$\mathbf{U} = u\mathbf{I} \tag{14}$$

a scalar matrix.

As was brought back into focus recently by Schönemann and Wang (1972) and McDonald (1972), most of what has been written about factor indeterminacy is thus equally relevant to classical test theory, where the problem has also been ignored in the past.

Factor Indeterminacy—A Numerical Example. Suppose a high school guidance counselor wants to recommend curriculum choices to four students, Smith, Jones, Todd, and Wilson, whose test scores are given in Table 1. The counselor accepts Spearman's theory, and he believes that a student's "general intelligence" should be the primary determinant of her curriculum choice. He performs a factor analysis to obtain the students' general intelligence factor scores. In this case, he might find the common factor pattern **a** and unique factor pattern **U** given in Table 1. However, when he attempts to determine **x**, the counselor encounters a serious problem. It turns out that, for the given **a** and **U**, there are many different "general ability" score vectors **x** which fit the factor model perfectly. This fact is referred to as "factor indeterminacy."

Table 1. Factor Indeterminacy—A Numerical Example

$$\begin{array}{cccccc} & \text{Jones} & \text{Smith} & \text{Todd} & \text{Wilson} & \\ \mathbf{Y} = & \left[\begin{array}{c} 1.36 \\ 1.54 \end{array}\right. & \begin{array}{c} .11 \\ -.35 \end{array} & \begin{array}{c} -1.47 \\ .03 \end{array} & \left.\begin{array}{c} .00 \\ -1.22 \end{array}\right] & \begin{array}{c} \text{Test 1} \\ \text{Test 2} \end{array} \end{array}$$

$$\mathbf{a} = \begin{bmatrix} .7071 \\ .7071 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} .7171 & 0 \\ 0 & .7071 \end{bmatrix}$$

$$\begin{array}{ccccc} & \text{Jones} & \text{Smith} & \text{Todd} & \text{Wilson} \\ \mathbf{x}_1' = & [1.12 & .85 & -.83 & -1.16] \\ \mathbf{x}_2' = & [1.60 & -1.09 & -.53 & .00] \end{array}$$

$$\begin{array}{cccccc} & \text{Jones} & \text{Smith} & \text{Todd} & \text{Wilson} & \\ \mathbf{T}_1 = & \left[\begin{array}{c} .79 \\ .79 \end{array}\right. & \begin{array}{c} .60 \\ .60 \end{array} & \begin{array}{c} -.59 \\ -.59 \end{array} & \left.\begin{array}{c} -.82 \\ -.82 \end{array}\right] & \begin{array}{c} \text{Test 1} \\ \text{Test 2} \end{array} \end{array}$$

$$\begin{array}{cccccc} & \text{Jones} & \text{Smith} & \text{Todd} & \text{Wilson} & \\ \mathbf{T}_2 = & \left[\begin{array}{c} 1.13 \\ 1.13 \end{array}\right. & \begin{array}{c} -.77 \\ -.77 \end{array} & \begin{array}{c} -.37 \\ -.37 \end{array} & \left.\begin{array}{c} .00 \\ .00 \end{array}\right] & \begin{array}{c} \text{Test 1} \\ \text{Test 2} \end{array} \end{array}$$

$$\begin{array}{ccccc} & \text{Jones} & \text{Smith} & \text{Todd} & \text{Wilson} \\ \hat{\mathbf{x}} = & [1.36 & -.12 & -.68 & -.58] \end{array}$$

For example, consider $\mathbf{x}_1$ and $\mathbf{x}_2$ in Table 1. As the reader can readily verify, both sets of general ability scores fit the factor model perfectly (within the limits of rounding error). However, $\mathbf{x}_1$ and $\mathbf{x}_2$ give radically different views of the intelligences of the students.

Smith, who would be labeled above average by $\mathbf{x}_1$, is far below average in $\mathbf{x}_2$. Wilson, rated far below average in $\mathbf{x}_1$, is exactly average in $\mathbf{x}_2$. The factor scores in $\mathbf{x}_1$ and $\mathbf{x}_2$ correlate only .33 with each other, and they are sufficiently disparate to have strongly contradictory implications. Again, it should be stressed that the data on hand, the test scores y_{ji}, provide absolutely no basis for distinguishing between these alternate solutions.

The identical problem exists in the test theory model. Algebraic substitution yields the contradictory sets of equally valid "true scores" $\mathbf{T}_1$ and $\mathbf{T}_2$, given in Table 1.

Does it then make any sense to say that Smith has "an intelligence score" or "a true score"? Can the factor model or the test theory model be of any use in prescribing a curriculum for Smith? Does either model tell us anything about Jones that we could not have concluded from the original test scores alone? These are the fundamental questions raised by factor indeterminacy.

A related problem is the meaning of "factor score estimation," or "true score estimation." As far back as the early 1930s, some factor analysts recommended "estimating factor scores" via a "regression approach." For the above example, the "regression estimates" of the factor scores would be $\hat{\mathbf{x}}$, as given in Table 1.

$\hat{\mathbf{x}}$ correlates equally well with $\mathbf{x}_1$ and $\mathbf{x}_2$, and the correlation is fairly high (.82). $\hat{\mathbf{x}}$ itself is not a solution, since it does not meet the constraints of the factor model. $\mathbf{x}_1$, a valid solution for $\mathbf{x}$, has declared Smith to be highly intelligent, whereas $\mathbf{x}_2$, an equally valid solution, has declared him subnormal. Does it alleviate the indeterminacy problem to "estimate" Smith's intelligence as "average," when such a judgment is based on estimated scores that do not fit the factor model at all? Guttman (1972) has commented incisively on these questions:

> Stating that $Y = T + E$, and that Y and all parameters are known for the bivariate normal distribution of T and E, leaves T very indeterminate for each individual. Many different score solutions will yield the same given bivariate parameters. This is an obvious mathematical fact. There is a known way for solving for all possible solutions for T (and E). Nothing in the statement of the original problem gives any way to choose one of these solutions as "the" solution.
>
> Instead of facing these mathematical facts, there has been some tradition of considering a different problem, namely: what is the regression of T on Y? Because regression equations involve only the bivariate parameters, and these are known, one doesn't actually have to pick out a particular solution for T in order to discuss the regression equation. Indeed, all possible solutions for T must have the same regression equation on Y, even though many of these solutions may be radically different from each other. The regression estimates turn out to be averages of all possible solutions. Again, using an average—which is *not* a solution—may be good for some purposes and bad for others. Defining a particular purpose may help pick out the particular exact solution required, obviating any thought of averaging or of regressions. The loss function implied by the regression of T on Y is quite different from loss functions involving choice of the wrong solution for T, and also from loss functions for using the regression estimate in place of T.

It appears that factor indeterminacy is a relatively basic problem of the factor model. It raises questions that, if left unanswered, might seriously compromise the ultimate purpose of the factor model.

Nevertheless, some evidence suggests that the problem is a relatively new one. The most widely cited historical summary of the early factor analysis literature, by Wolfle (1940), which includes a major section on the "limitations of factor analysis," makes no mention whatever of factor indeterminacy. From 1940 to 1951, not a single article on the subject was published. Most of the more popular texts, such as Thurstone (1947) and Harman (1960), omit any reference to factor indeterminacy. A mere handful of articles have appeared since 1955, most of them since 1970. Thus, readers who have to rely on the more popular sources of information might well conclude that the indeterminacy issue is a very recent development and is only now gaining momentum.

In fact, however, factor indeterminacy is not a new issue at all. Indeed, it is almost as old as factor analysis itself. Spearman published "The Abilities of Man" in 1927. More than 15 articles on indeterminacy were published in the succeeding 10 years, beginning with Wilson (1928a). Many of the major technical facts of indeterminacy were developed in these papers. The significance and interpretation of these facts were debated at length. In several instances, these debates became rather intense, and the articles still make lively and interesting reading.

The history of factor indeterminacy is thus rather uneven. Periods of great activity have alternated with periods of almost total neglect. Some writers have attached great significance to the issue. Others have dismissed it as trivial. Most have completely ignored it. Since the issue is again attracting considerable attention, this is an excellent time to reexamine its history and to take stock of 49 years of progress. The main objectives of the present account are therefore (1) to provide the reader with a relatively thorough but nontechnical review of the factor indeterminacy issue; (2) to fill in the gaps in Wolfle's historical record; and (3) to reassess the whole trend of the events of the 1930s in light of current developments in psychometrics. There is a definite parallel between the way factor indeterminacy was treated in the 1930s and the way it is being treated in the 1970s. Discussing this parallel, we hope, will deepen our understanding not only of factor indeterminacy but of factor analysis itself.

FACTOR ANALYSIS BEFORE 1928

> The average intercolumnar correlation from the tables of 14 different investigators, summarizing 30 years of psychological researches and representing a

> great wealth of test material, was unity, as expected by the unifocal hypothesis of a general factor. It seemed to be the most striking quantitative fact in the history of psychology [Dodd, 1928a, p. 214].

Although historians generally concentrate on Charles Spearman's pioneering work in correlation and factor analysis, he should also be recognized as one of the first mathematical psychologists. Spearman proposed two-factor theory in 1904 as a mathematical, empirically falsifiable, psychological model. He and his associates then launched a program of experimental verification that was revolutionary in scope and rigor. Thus, when Hart and Spearman concluded in 1912 that two-factor theory had indeed been verified, they had just cause for enthusiasm. An impressive amount of empirical evidence seemed to weigh overwhelmingly in the theory's favor.

Hart and Spearman envisioned their factor model, especially factor scores, as the central focus of a new educational technology, from which would ensue a wide range of practical benefits:

> Indeed, so many possibilities suggest themselves that it is difficult to speak freely without seeming extravagant.... It seems even possible to anticipate the day when there will be yearly official registration of the "intellective index," as we will call it, of every child throughout the kingdom.... The present difficulties of picking out the abler children for more advanced education, and the "mentally defective" children for less advanced, would vanish in the solution of the more general problem of adapting education to all.... Citizens, instead of choosing their career at almost blind hazard, will undertake just the professions really suited to their capacities. One can even conceive the establishment of a minimum index to qualify for parliamentary vote, and above all for the right to have offspring [Hart and Spearman, 1912, pp. 78–79].

Godfrey H. Thomson's sampling theory of abilities provided two-factor theory with its most serious challenge during this early period. Thomson's theory asserts that the mind is composed of an extremely large number of components and that some of these (higher-level units) participate in many different kinds of activities, while others (lower-level units) are restricted to a single kind of activity. Any given task is performed using a random sample from the appropriate units of both levels. Thomson accepted Spearman's data, but insisted (1916, 1919) that they agreed with his own theory as well as Spearman's.

At this time, debate focused on whether the tetrad difference criterion sufficed to prove the actual "existence" of **g**, Spearman's general ability factor. Spearman was not always careful to acknowledge the distinction between proving the *compatibility* of data and mathematical system and proving the empirical *existence* of the mathematical system's constructs. (Thomson commented extensively on this point in 1935, and Wolfle, p. 7, followed suit in 1940.) Spearman concentrated on showing that the "hierarchy" of the system of correlations implied that the data were compatible with factor theory.

Garnett (1919) gave a proof that, in the case of normally distributed variables, the Spearman theory must necessarily be compatible with a correlation matrix exhibiting hierarchy. A year later, Garnett (1920) compared Spearman's theory with Thomson's and expressed a preference for the former on the grounds of parsimony. Garnett also gave an erroneous proof that when conditions of hierarchy are satisfied, **g** is "uniquely determined."

By 1922, Spearman was able to dispense with Garnett's normality assumption. He demonstrated that, for an infinite number of tests, "hierarchy" implied that the two-factor theory would fit the data. More important, he gave a proof covering the situation where the number of tests is not large. Spearman gave a determinantal expression for **g**, which expressed it as a function of two components, one of which is a determinate linear combination of the original tests, the other a function of a variable **i**, which could be "any new variable uncorrelated with all the others."

Spearman's aim was to prove the *existence* of **g**, which for him was synonymous with showing the existence (under condition of hierarchy) of a set of factor scores fitting the single factor model. Unwittingly, he also demonstrated the indeterminacy of **g**, since more than one satisfactory **i** existed.

In the years immediately following, Spearman remained unaware of the full implications of **i**. In 1927, he repeated his 1922 proof almost verbatim in the mathematical appendix to "The Abilities of Man" and gave particular emphasis to Garnett's erroneous proof of the "uniqueness" of **g**: "There is another particularly important limitation to the divisibility of the variables into factors. It is that the division into general and specific factors all mutually independent can be effected *in one way only;* in other words, it is unique. For the proof of this momentous theorem, we have to thank Garnett" (Spearman, 1927, p. vii). Ironically, this enthusiastic declaration of the "uniqueness" of **g** occurred just two pages after the derivation revealing the indeterminate part **i**.

Although Spearman was still unaware of the indeterminacy of **g**, he showed keen interest in its *linear unpredictability* from the original tests in his Section 4, pages xvii–xviii, titled "To Measure a Person's **g**." (The important distinction between the *indeterminacy* of **g** and its *linear unpredictability* from the original tests will be discussed later in this chapter.)

The Spearman-Thomson debate continued through the 1920s. Dodd (1928a, 1928b) chronicled this controversy, as well as the other significant developments of the period, in a thorough and well-organized review. The parsimony of two-factor theory had apparently tipped the balance of public opinion in its favor by this time. Dodd, expressing what then was probably the prevailing view, concludes, "For the purpose of measurement and prediction, the concepts of **g** and **s** are the more useful" (1928b, p. 278).

Since Dodd wrote his reviews just before Wilson's first paper on factor indeterminacy, he never commented on the issue. Like Spearman, Dodd failed to see the full significance of the arbitrary variable **i**. Nevertheless, his papers provide an excellent summary of the "preindeterminacy" era.

E. B. WILSON AND THE EARLY FOUNDATIONS

> If any event is more likely than another to quicken the progress of psychological mathematics, it is the entry on the scene of a mathematician so eminent and so free from prejudice as Professor E. B. Wilson [Spearman, 1929, p. 212].
>
> The development of factor theory, as well as its applications in science, will be accelerated by the assistance of mathematicians; and it is gratifying that Professor E. B. Wilson has turned his attention to these problems in several papers. The future development of factor analysis in psychology will probably require more mathematical competence than we can supply in our own ranks [Thurstone, 1935, p. xi].

E. B. Wilson, virtually an unknown figure among factor theorists today, was one of America's premier mathematicians in the 1920s. He wrote the standard texts of the period on both vector analysis and advanced calculus and was president of the American Statistical Association in 1929. As the foregoing quotations show, both Spearman and Thurstone were quick to acknowledge his stature.

In 1928, reviewing "The Abilities of Man" for *Science,* Wilson pointed out that *different* sets of factor scores could fit Spearman's model equally well, for the same set of data. Wilson's simple treatment (1928a) is noteworthy in that, besides pointing out indeterminacy explicitly for the first time, it (1) provides the first numerical illustration of indeterminacy and (2) provides the first attempt to numerically characterize the extent of factor indeterminacy. The numerical example is a factor analysis of the scores of six students on three tests. Wilson derives the extreme alternate values for the "general intelligence" factor scores for each student and uses the difference between these extreme values in characterizing the extent of indeterminacy.

Wilson (1928b) formalized his conceptions of the indeterminacy of Spearman's model with a geometric description of its implications for the nature of **g**. Wilson stated in his equation

$$\mathbf{g} = \mathbf{v} + \mathbf{r}_o \tag{15}$$

that factor **g** can be expressed as the sum of two components, one a determinate linear combination of the original variables, the other an indeterminate, largely arbitrary, component. Thus Wilson clarified with his geometric presentation what was already inherent but unnoticed in Spearman's algebra (1922, 1927).

Following a third Wilson article (1929a) on factor analysis, a review of T. L. Kelley's "Crossroads in the Mind of Man: A Study of Differentiable Mental Abilities," Spearman (1929) published a brief article in defense of the two-factor theory. Although Wilson portrayed factor indeterminacy as a lack of uniqueness in the variable **g**, Spearman presented the issue in a different light. Spearman admitted that **g** was not uniquely defined and protested that he had been "urging very much the same thing" (in Spearman, 1927, pages xvii–xviii. Actually, however, these were the pages where Spearman discussed the linear unpredictability of **g** from the original tests). Spearman suggested that the indeterminacy of **g** could be eliminated by introducing a new test into the test battery, a test constructed to correlate perfectly with **g** itself. He indicated (1929, p. 214) that "unpublished research in our laboratory has more than once obtained for an r_{ag} values of .99," adding that "nothing stands essentially in the way of raising it much higher still; in fact as near as desired to unity."

In a rejoinder, Wilson (1929b) pointed out that if a test **a** correlates perfectly with **g**, we can "throw away our scaffolding," that is, forget about factor analysis entirely and simply make test **a** the measure of **g**. Had his perspective on the problem been more advanced, Wilson might also have added that since there was more than one **g**, there would be more than one such **a**, and so adding a particular **a** would be tantamount to arbitrarily selecting a particular **g**. Spearman apparently accepted Wilson's rebuttal and did not publish another paper on indeterminacy for another four years.

Wilson contributed impressively to the foundations of factor analytic theory. His early papers on factor indeterminacy introduced many of the significant ideas developed and extended by other writers in the 1930s. In his later papers on factor analysis, Wilson pointed out other problem areas, such as the "identifiability" problem. Clearly, his recognition as one of the key figures in the history of factor theory is long overdue.

THE DEBATES OF THE 1930S

> The particular indeterminateness of **g**-measurements ... is nothing else than the error just mentioned as being due to the limited number of tests available for the purpose of measuring. Moreover, it is nothing more than the probable error given by Holzinger for the regression equation [Spearman, 1933, p. 108].

> It is tedious to drag out a controversy after the scientific evidence is once agreed on and published where everyone may examine it for himself, but I cannot but ask a brief space in order to point out as tersely as possible and without adjacent forbidding-looking mathematics two quite definite errors which it seems to me Spearman has made [Camp, 1934, p. 260].

Technical Developments. A number of writers in the 1930s added to Wilson's foundation in further developing the major facts of indeterminacy. These technical developments can be summarized under three major themes: the construction approach, indeterminacy in the limit, and the transformation approach.

The construction approach. H. T. H. Piaggio published several articles on factor indeterminacy and produced a number of theoretical innovations of major importance. In the first article, in 1931, he simplified Spearman's complicated determinantal formula for **g** (1922) and showed explicitly that **g** can be divided into a determinate and an indeterminate part. Irwin (1932) pointed out the close relationship between Wilson (1928b) and Piaggio (1931). In 1933, Piaggio gave a much more explicit statement and proof of his 1931 result, which had since been verified by Heywood's independent derivation (1931). He showed that **g** can be expressed as

$$\mathbf{g}' = \mathbf{a}'\mathbf{G}_{yy}^{-1}\mathbf{Y} + p\mathbf{s}' = \hat{\mathbf{g}}' + \mathbf{e}' \tag{16}$$

where **s** is an arbitrary vector of numbers in standard score form meeting the constraint that $\mathbf{Ys} = \mathbf{0}$, and p is a scalar defined as $p = (1 - \mathbf{a}'\mathbf{C}_{yy}^{-1}\mathbf{a})^{1/2}$. This equation provides a means of constructing different sets of common factor scores for any set of data **Y** and factor pattern **a** by choosing different values for the arbitrary vector **s**. This approach to indeterminacy became known later as the "construction approach." Piaggio (1933) demonstrated the sufficiency of Equation (16). That is, any set of numbers **g** constructed by Equation (16) would fit the factor model given in Equations (5) to (7). Furthermore, in 1935, Piaggio demonstrated the necessity of Equation (16), thus establishing that any and all sets of factor scores satisfying the factor model must be expressible in terms of Equation (16).

Indeterminacy in the limit. If the two-factor theory holds, then the size of the indeterminate part of **g** becomes infinitesimally small as the number of variables in the study becomes infinitely large. This was first pointed out by Spearman (1922) and Piaggio 1931) and clarified in Piaggio and Dallas (1934) and Irwin (1935). Thus, if a single common factor can explain an infinite number of variables, the indeterminacy will vanish in the limit.

The transformation approach. Thomson (1935) introduced an approach to indeterminacy which Schönemann and Wang (1972) later called the "transformation approach." Suppose the factor model, for a given **a**, **U**, is written in the form

$$\mathbf{Y} = \mathbf{a}\mathbf{x}' + \mathbf{UZ} = [\mathbf{a}:\mathbf{U}] \begin{bmatrix} \mathbf{x}' \\ \cdot\cdot \\ \mathbf{Z} \end{bmatrix} \tag{17}$$

Thomson (1935) gave the formula for a transformation matrix **B** with the properties

$$\mathbf{BB}' = \mathbf{I} \tag{18}$$

and

$$[\mathbf{a}:\mathbf{U}]\mathbf{B} = [\mathbf{a}:\mathbf{U}] \tag{19}$$

The existence of **B** satisfying Equations (17) to (19) implies factor indeterminacy, since we can then write

$$\mathbf{Y} = [\mathbf{a}:\mathbf{U}]\mathbf{B}\begin{bmatrix}\mathbf{x}' \\ \cdot\cdot \\ \mathbf{Z}\end{bmatrix} = [\mathbf{a}:\mathbf{U}]\begin{bmatrix}\mathbf{x}^{*\prime} \\ \cdot\cdot \\ \mathbf{Z}^{*}\end{bmatrix} \tag{20}$$

Thus, for any factor scores **x** and **Z** satisfying the factor model, there exists also a different set of factor scores $\mathbf{x}^{*\prime}$, $\mathbf{Z}^{*}$, which also satisfy the model.

Ledermann (1938) extended Thomson's result to give a formula for **B** in the multiple factor case.

Interpretation. The technical facts of indeterminacy were established with little debate, but there were major differences of opinion over the interpretation of these facts. We shall summarize some key areas of disagreement which dichotomized the views of those who felt that indeterminacy was a serious problem and those who felt it was not.

Indeterminacy—lack of uniqueness or error of measurement? Wilson characterized factor indeterminacy as a lack of uniqueness. The fact that an infinite number of different sets of factor scores all satisfied the factor model meant that the model's central concept **g** was not uniquely defined. He saw this lack of uniqueness as seriously compromising the practical value of the model.

Camp (1932, 1934) shared this view and expressed it rather forcefully: "If, before looking at Smith's scores on the tests, one may choose a number at random (subject only to the broad limitations mentioned before), and can then demonstrate that this number can be assigned as Smith's **g**, as well as any other number, and in perfect harmony with all the other hypotheses, then it is meaningless to assert that Smith has a **g**" (Camp, 1934, p. 261).

Spearman (1933, 1934), in trying to defend his two-factor theory, portrayed factor indeterminacy as essentially a sampling problem, nothing more than an error of regression estimate. Spearman's "regression analogy" argument was prompted by the formal similarities between the algebra of multiple regression and Piaggio's construction equation. As noted previously, any and all factor solutions can be expressed

$$\mathbf{g}' = \mathbf{a}'\mathbf{C}_{yy}^{-1}\mathbf{Y} + p\mathbf{s}' = \mathbf{b}'\mathbf{Y} + p\mathbf{s}' \tag{21}$$

$$= \mathbf{b}'\mathbf{Y} + \mathbf{e}' = \hat{\mathbf{g}}' + \mathbf{e}' \tag{22}$$

In multiple regression, the object is to choose the linear combination of predictor variables **Y** that best predicts a known criterion **g**. Here too we may write

$$\mathbf{g}' = \mathbf{b}'\mathbf{Y} + \mathbf{e}' = \hat{\mathbf{g}}' + \mathbf{e}'$$

Thus, the identical equation can be used to describe the factor construction and multiple regression situations. Spearman emphasized this fact in declaring indeterminacy to be nothing more than the probable error of the regression equation. This probable error could be reduced by simply adding more tests to the test battery. (Although it may not be immediately obvious, this argument was essentially a restatement of Spearman's earlier (1929) position. Instead of adding one test to improve the determinacy of **g**, Spearman now proposed to add many.)

In rebuttal, Thomson (1934) pointed out that although the regression and construction situations have strong algebraic similarities, there is a subtle but crucial difference between them.

> The comparison shows clearly that the indeterminate term ... is the same kind of thing in Spearman's eq ... as in ... the ordinary use of the regression equation, *except for the important difference that in Spearman's case there is no measure of* **g** *other than that arising from the team of tests.*
>
> In the prediction, by tests at entrance, of marks in the senior year at a university, there is no doubt about the actual existence of these latter. They are awarded by quite independent means and the accuracy of prediction can be checked. ... In the Spearman case, however, there is no other evidence of the existence of **g**

> other than the magnitude predicted for it. It arises out of the team of tests, and is not measured in any other way than by the team of tests. Its only attributes are mathematical estimates, and to the extent to which these are indeterminate, one may perhaps hold that **g** itself, being nothing else, is indeterminate.
>
> This distinction between the two cases may appear to be subtle, but it seems a proper distinction to draw. It is of course convenient to make the hypothesis that a real **g**, perfectly determined, exists and that the quantity **i** expresses merely an uncertainty in the measurement, not any doubt as to its existence. But it is only a hypothesis in this case. In the other case there indubitably was something with an existence independent of the estimate [Thomson, 1934, pp. 96–97].

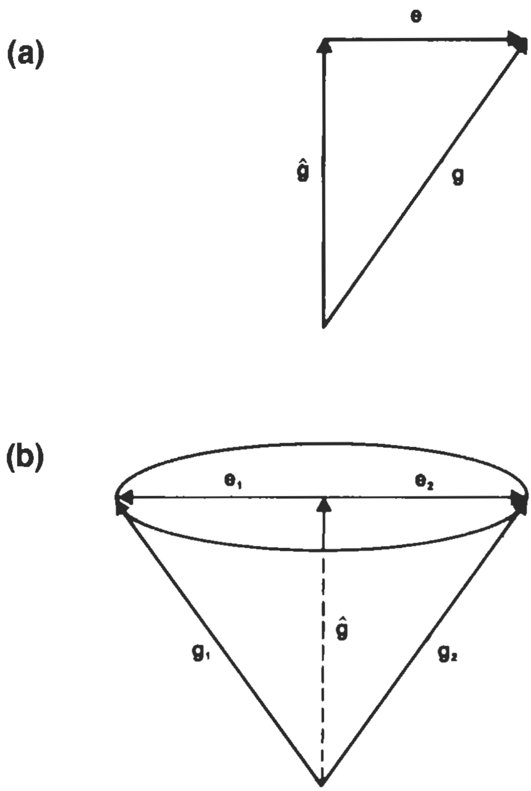
In multiple regression we attempt to *predict* a criterion from a set of variables. The criterion is always known and in principle directly measurable, although decisions made on the basis of our predictions may restrict our ability to measure the criterion.

In factor analysis, the regression equation serves as a *defining model* for a latent variable that is never directly measurable. Any criterion variable that fits the regression equation fits the factor model and is, by definition, an alternative version of a factor.

To help clarify this distinction, we briefly consider the geometry of factor indeterminacy and multiple regression. In multiple regression, the regression weights define a best linear estimate $\hat{\mathbf{g}}$, for predicting the criterion $\mathbf{g}$ from the original tests. $\mathbf{g}$, $\hat{\mathbf{g}}$, and $\mathbf{e}$ are all directly measurable (in principle) and are uniquely defined. If $\mathbf{g}$, $\hat{\mathbf{g}}$, and $\mathbf{e}$ are represented as vectors, we obtain the geometric representation in Figure 1(a).

In Spearman's model, we deal with a latent variable that is not directly measurable *and that is defined only in terms of its ability to satisfy Equations (5) to (7).* An infinity of variables satisfies these equations. They all fit a particular regression equation; that is, they all have common variance-covariance properties with the original tests. Hence, in factor analysis, the regression equation serves as a *model* for the variable $\mathbf{g}$. Any variable that satisfies

Figure 1. Geometric Comparison of Multiple Regression (a) and Factor Construction (b)



the constraints of Equation (16) will fit the factor model. These variables are represented in Figure 1(b). They lie on a (hyper) cone circumscribed about $\hat{\mathbf{g}}$. In the factor analysis situation, we have not only a linear unpredictability but also an *indeterminacy*. The factor model seeks to define $p + 1$ variables uniquely by means of p equations. For finite p, this is impossible.

Thomson's argument went largely unheeded. In 1935, Piaggio suggested using the determinate parts of the general and specific factors as "approximate" factors. This, in essence, solves the indeterminacy problem by ignoring it. Piaggio did not draw attention to the fact that such "approximate" factors could not possibly fit the factor model. Instead, he concentrated on the virtues of getting "rid of" the indeterminacy problem.

The measurement of indeterminacy. Advocates of the lack of uniqueness position stressed the conceptual difficulties inherent in the existence of disparate, conflicting solutions for **g**. The greater the possible conflict in empirically indistinguishable factor scores, the greater the indeterminacy problem, since it was difficult to conceive of a meaningful concept of **g** which could alternatively assign a person vastly different scores. In keeping with this view, Wilson (1928a) and Camp (1932) quite naturally referred to sets of widely different solutions in characterizing the extent of factor indeterminacy.

Spearman and Piaggio, proponents of the "regression analogy," characterized indeterminacy as the linear unpredictability of the factor from the original variables. They used variants on the multiple correlation coefficient between **g** and the original tests as numerical indices of indeterminacy. Along these lines, Piaggio (1933) offered two indices. One is the ratio of the standard deviations of the indeterminate and determinate parts of **g**. The second, the square of the first, is the ratio of indeterminate to determinate variances. The second ratio could be preferred on the mathematical grounds that the determinate and indeterminate variances always sum to unity and are hence in a consistent metric. Moreover, since both ratios are usually less than 1, the second index almost inevitably yields substantially smaller numbers than the first. Spearman (1933, p. 108) therefore expressed a decided preference for the latter index:

> The work of Piaggio would seem to paint the case in still darker colours. Now, what he writes seems to me to be unimpeachably accurate so far as it goes. But I do think that in his first method of estimating the degree of indeterminacy he has indicated the most unfavorable viewpoint. Far better would appear to be his second method, which is that of Professor Holzinger. Here the natural and most significant basis for comparing the relative influences of the determinate and the indeterminate parts of the measurement of **g** is

> taken to be, not their respective standard deviations, but rather their variances. Upon adopting this improved method of comparison, the indeterminate influence makes a startling drop; for instance, from 10 to only 1 percent.

Thus, methods for numerically characterizing the extent of indeterminacy were intimately related to the theoretical conception of indeterminacy itself. This trend has persisted into the 1970s.

The indeterminacy issue lost momentum as the 1930s drew to a close. Wilson concentrated on a different series of problems, including the "identifiability" of the factor model's variance-covariance parameters. Other writers remained isolated with opposing theoretical positions. When Wolfle (1940) wrote his historical review, factor indeterminacy was not longer attracting the attention in the literature that it had been just a few years before. Nevertheless, it was without a doubt one of the most significant theoretical issues discussed in the period from 1928 to 1939. Wolfle's failure to ever mention the indeterminacy issue in his paper seriously impairs his stature as an objective historian of "Factor Analysis to 1940."

THE THURSTONIAN ERA: 1940–1951

> If the scientist takes his numerical coefficients very seriously at the exploratory stage, he may be lacking in a desirable sense of humor about the crudeness of all his tools in spite of their polished appearance. Too much concern with numerical minutiae at that stage may lead him astray from the conceptual formulations that constitute his real goal [Thurstone, 1947, p. xi].

Factor analytic literature became increasingly atheoretical in the 1940s. Thurstone and his associates, whose efforts established the United States as a new center of research in the field, invested remarkable energy in the achievement of technical refinements. However, they generally ignored such basic theoretical problem areas as factor indeterminacy, identifiability of the factor pattern, and invariance of factors under linear transformations of the original variables. Thus, the period produced primarily statistical and computational advances, perhaps the most significant of which were (1) the popularization of "multiple factor analysis," (2) the introduction and widespread acceptance of the simple structure concept as a solution of the "rotation problem," and (3) Lawley's investigations of the statistical aspects of factor analysis, culminating in his maximum likelihood solution.

The multiple factor model, a straightforward generalization of Spearman's single factor case, is often attributed to Thurstone (for example, by McNemar, 1951, and, indeed, by Thurstone himself, 1947). Actually it is due to Garnett (1919). Dodd (1928a) discussed Garnett's contribution in detail, while describing multiple factor analysis as "a tool of possibly immense value for the quantitative analysis of psychological data" (p. 226). Curiously, the article immediately following Dodd's in the *Psychological Review* was written by L. L. Thurstone.

In the multiple factor model, Equations (4) and (5) are generalized to

$$ {}_n\mathbf{Y}_N = {}_n\mathbf{A}_p\mathbf{X}_N + {}_n\mathbf{U}_n\mathbf{Z}_N \tag{23}$$

where

$$\mathbf{XZ}' = \mathbf{0} \tag{24}$$

and

$$\frac{\mathbf{ZZ}'}{N} = \mathbf{I}_n \tag{25}$$

If one defines $\mathbf{C}_{yy} = \mathbf{YY}'/N$ and $\mathbf{C}_{xx} = \mathbf{XX}'/N$, then

$$\mathbf{C}_{yy} = \frac{1}{N}(\mathbf{AX} + \mathbf{UZ})(\mathbf{AX} + \mathbf{UZ})' = \mathbf{AC}_{xx}\mathbf{A}' + \mathbf{U}^2 \tag{26}$$

The multiple factor model has a "rotation problem" not present in the single factor case, since $\mathbf{Y} = \mathbf{AX} + \mathbf{UZ}$ implies $\mathbf{Y} = \mathbf{ALL}^{-1}\mathbf{X} + \mathbf{UZ} = \mathbf{A^*X^*} + \mathbf{UZ}$, where $\mathbf{A}^* = \mathbf{AL}$ and $\mathbf{X}^* = \mathbf{L}^{-1}\mathbf{X}$ for any nonsingular $\mathbf{L}$. Thus, for any given $\mathbf{A}$ and $\mathbf{X}$ that satisfy Equation (23), one may choose an infinite number of other **A*** and **X***, which also satisfy it. Thurstone proposed his "simple structure" criterion to resolve this rotational indeterminacy. Simple structure is based on the idea that the most useful and readily interpretable solution is the most parsimonious one, that is, where each test involves the smallest number of common factors. It added an air of objectivity to the choice of a factor pattern and was readily adaptable to computers. The criterion has enjoyed widespread popularity.

The growing interest of statisticians in the factor model yielded some significant advances. Beginning in 1940, Lawley published a sequence of papers in which he gave equations for maximum likelihood estimates of factor loadings and developed a basis for statistical testing in factor analysis. However, computational difficulties delayed the practical implementation of these theoretical results until subsequent advances in computer technology and numerical analysis occurred (Howe, 1955; Jöreskog, 1967).

The computational and statistical advances of the 1940s were impressive, especially when one considers the disruptive influence of World War II. Somewhat less impressive was the spreading amnesia that engulfed the psychometric community concerning the perplexing flaws in the factor model's algebraic structure. No articles on factor indeterminacy appeared during this decade. The major texts of the period (Thurstone 1947; Holzinger and Harman, 1941) followed Wolfle's precedent in simply ignoring factor indeterminacy, together with the other theoretical problems Wilson had uncovered.

ERA OF BLIND FACTOR ANALYSIS: 1952–1969

> In this period factor analysis was frequently applied agnostically as regards structural theory to all sorts of data.... The hope was that factor analysis could bring order and meaning to the many relationships between variables [Mulaik, 1972, p. 9].

The pragmatic trend of the 1940s continued in the two subsequent decades, which Mulaik (1972) characterized as an "era of blind factor analysis." Factor analytic methodology became increasingly refined, and computer technology rendered its application virtually effortless. The technique became accessible to a broad spectrum of users, not all of whom were familiar with its theoretical underpinnings. Consequently, factor analysis was used more and more as a data reduction technique, rather than a model.

A practical problem that received a great deal of attention in the 1950s and 1960s was the "estimation" of factor scores. Piaggio (1933) had shown how to construct "factor scores" for the single factor case, but factor analysts of the early 1950s were apparently unaware of this work. A common cliché of the period, dating back at least to Thomson (1948), was that "factor scores cannot be computed, they can only be estimated." In 1952, Kestelman paved the way for the generalization of Piaggio's result to the multiple factor case, by proving that if **Y**, **A**, and **U** satisfy Equation (26), then a matrix **X** of factor scores satisfying Equation (23) can always be constructed.

Kestelman (1952, p. 2) alluded to factor indeterminacy only briefly: "Indeed, the distinction between unique and exact values is often overlooked, and it is therefore desirable to emphasize at the outset that, even when the uncorrelated values are exact . . . , they need not necessarily be unique." He also remained noncommittal about the status computable factor scores should assume: "Even when we assume that the factors are more numerous than the tests, exact numerical specifications can still be found for each of the factors, such that the factor measurements obtained will be in standard measure and entirely uncorrelated. Of course, it does not follow that such measurements will necessarily be superior from a statistical standpoint. There may be theoretical or practical grounds for preferring the correlated estimates calculated by the equations in current use (for example, the fact that they have been so determined as to give the 'best' fit, judged by the principle of least squares). But these are further issues with which this article is not concerned" (p. 2).

Guttman (1955) produced several important theoretical results and, in contrast to Kestelman, commented unequivocally on their significance. Some of Guttman's more important results are the following:

1. He proved that for oblique as well as orthogonal factors, and for the population as well as the sample, if **Y**, **A**, and **U** satisfy Equation (26), then an **X** exists satisfying Equation (23). This extended Kestelman's result, which was restricted to orthogonal factors in the sample.

2. He gave a generalized construction formula, applicable to both sample and population cases and orthogonal and oblique factors. He also proved its necessity and sufficiency, thus greatly extending the earlier result of Piaggio (1935). In the case of orthogonal factors, **X** satisfying Equation (15) can always be constructed as

$$\mathbf{X} = \mathbf{A}'\mathbf{C}_{yy}^{-1}\mathbf{Y} + \mathbf{PS} \tag{27}$$

$$\mathbf{PP}' = \mathbf{I} - \mathbf{A}'\mathbf{C}_{yy}^{-1}\mathbf{A}, \tag{28}$$

$$\mathbf{SS}' = \mathbf{I} \qquad \mathbf{YS}' = \mathbf{0} \tag{29}$$

$$\mathbf{SJ} = \mathbf{0} \tag{30}$$

with **J** a conformable vector of 1s. Piaggio's result is an obvious special case of Equation (27).

3. He introduced the minimum correlation between alternate factors as a numerical index of indeterminacy. Guttman examined the relation between this correlation, which he called ρ^*, and ρ, the multiple correlation between the factor and the observed variables. The relation is

$$\rho^* = 2\rho^2 - 1 \tag{31}$$

He commented on some data where ρ varies from .630 to .908, and ρ^* thus varies from $-.206$ to .649: "It seems that the sought-for traits are not very distinguishable from radically different possible alternative traits for the identical factor loadings" (Guttman, 1955, p. 74).

4. He questioned the usefulness of the concept of second order factoring, since such factors would generally be highly indeterminate.

5. He greatly extended Piaggio's result (1931) by estab-

lishing conditions under which indeterminacy vanishes in the limit in the multiple factor and population cases.

Guttman pointed out that the common practice of naming factors according to the variables that have high loadings on them was illogical, when for any set of loadings there exists an infinite number of different factor solutions. He concluded, "the Spearman-Thurstone approach may have to be discarded for lack of determinacy of its factor scores" (1955, p. 79).

Guttman's negative conclusions failed to strike a responsive chord within the American psychometric community, partly, perhaps, because they were published in a relatively inaccessible British journal. The technical development of factor analysis continued to dominate the 1960s, during which factor indeterminacy and other significant theoretical issues received scant attention.

Heermann (1964, 1966) summarized the previous results on the geometry and algebra of factor indeterminacy in two very readable reviews. Although Heermann presented a relatively coherent account of the facts of indeterminacy, his conclusions on the meaning of it all lacked focus and were on occasion self-contradictory. He remarked, "Without doubt, the generality and scientific utility of the factor model would be enhanced if some meaningful method could be found to render factor scores determinate" (1964, p. 380). However, he rejected its most popular determinate alternative, component analysis, because components "are always contained in the test space, and cannot be expected to represent anything which goes beyond the original measures" (p. 380). He does not mention that the increment (the term **PS** in Equation 27) by which factor scores go "beyond the original measures" is arbitrary and thus hardly can be counted on to provide information about the state of nature.

Heerman concludes that "factor analysis does not seem very useful in describing the individual subject," but that it still may be useful for the study of covariance structures, since "description of the individual subject is not necessarily a major objective of factor analysis" (1964, p. 379). The widespread popularity of this view was apparent in the cursory and rather oblique treatment afforded factor scores in such texts as Thurstone (1947) and Harman (1960, 1967).

Several writers continued to explore aspects of "factor score estimation"; McDonald and Burr (1967) and Harris (1967) investigated the properties of several such estimates. Harris discussed factor indeterminacy at some length and justified factor score estimation by noting that "it is well known that 'true' factor scores are not uniquely computable, and thus in one practical sense are of no use at all." The treatment of the indeterminacy problem by McDonald and Burr is quite brief. They note perfunctorily that "the theoretical problems connected with the indeterminacy of factor scores (Guttman, 1955) have led to some misgivings about the usefulness of the common factor model." They neither describe the misgivings nor attempt to dispel them. They then repeat the cliché that "factor scores cannot be determined precisely, but only 'estimated,'" without answering Guttman's question about how such estimation can be meaningful when the factors themselves are not uniquely defined.

Thus, as the 1960s drew to a close, factor indeterminacy was still a relatively quiescent issue. Commenting on the public reaction to the "existence, properties, and consequences" of factor indeterminacy, Heermann (1964) noted that there was "some confusion." He might well have added "considerable complacency." By 1970, technology had advanced to the point where a 100 × 100 correlation matrix could be factor analyzed in just a few minutes. An immense and impressive superstructure had replaced the mere "scaffolding" of Wilson's day, and yet the foundations of the edifice remained largely untested.

FACTOR INDETERMINACY FROM 1970 TO 1976

> The theoretical scores are not available: consequently, several systems for estimating the scores have been proposed [Tucker, 1971, p. 427].

After three decades of relative inactivity, factor indeterminacy has again become the subject of debate. More articles on the subject have appeared from 1970 to 1976 than in the previous 30 years combined. Although indeterminacy has hardly become a "popular" issue, sections on the problem in recent texts (Mulaik, 1972; Gorsuch, 1974) appear to reflect an increased awareness of its existence.

Schönemann (1971) derives a simplified formula for the minimum average correlation between equivalent sets of (orthogonal) common and unique factors. He begins by noting the lengthy and somewhat obscured historical antecedents of the indeterminacy issue. He then offers a simplified proof of Ledermann's result. This approach yields a generalized representation for the orthogonal right unit matrix **B** in Equation (19), which in turn allows the derivation of an upper bound for the minimum average correlation between equivalent sets of uncorrelated factors, given by

$$\bar{r}_{\min} = \frac{p - m}{p + m} = \frac{1 - (m/p)}{1 + (m/p)} \qquad (32)$$

where p is the number of original variables, and m is the number of factors extracted. This result shows that, for given m and p, the indeterminacy index is independent of the data, a point not noticed by previous authors.

Schönemann also offers a proof that all equivalent factors are related by a Ledermann transformation. This proof contains an error (Schönemann, 1973) and in fact only applies to all linearly related equivalent factors. However, this error does not affect the substantive conclusions of the paper, which are based on an upper bound for the minimum correlation, since the existence of additional factors not related by a Ledermann transformation could not raise this upper bound but could only lower it still further.

The Schönemann (1971) results are limited to uncorrelated factors, and the $\bar{r}_{\min}$ statistic includes both common and unique factors, whereas the common factors are of primary interest in practice.

Meyer (1973), using results from Guttman (1955, 1956), derived an equation relating the average indeterminacy of common and unique factors to the ratio of the number of factors to the number of variables.

Schönemann and Wang (1972) extended the Schönemann (1971) results in an investigation of both empirical and theoretical aspects of the relationship between indeterminacy and maximum likelihood factor analysis (MLFA), which, largely through the efforts of K. Jöreskog, had by this time become the preferred method for applying the factor model in practical work. The paper yielded a number of new results:

1. If **A** and $\mathbf{U}^2$ are identified by the diagonality constraint

$$\mathbf{A}'\mathbf{U}^{-2}\mathbf{A} = \text{diagonal} \tag{33}$$

then the uncorrelated common factors associated with **A** are ordered from most to least determinate among all common factors obtainable by orthogonal or oblique rotation.

2. The measure of indeterminacy for these factors is a simple function of the latent roots of an eigenproblem that is routinely solved in the course of an MLFA. Thus, the factor analyst using MLFA can simultaneously assess goodness of fit and factor indeterminacy in an easy and convenient way.

3. Factor scores fitting the factor model can be computed from the maximum likelihood estimates of **A**, $\mathbf{U}^2$, and **Y**, whether $\mathbf{C}_{yy}$ fits the factor model or not.

4. Despite the long history of the indeterminacy problem, there had been almost no data assessing the extent of the indeterminacy that the factor analyst could expect to encounter in practice. To provide some evidence on this point, Schönemann and Wang analyzed data from 13 different factor analytic studies, using MLFA and computing measures of factor indeterminacy and goodness of fit. Some major findings emerged from this wide range of empirical data. First, as the number of variables in a study increases, the number of factors required to achieve an adequate statistical fit to the factor model also increases. On the other hand, increasing the number of factors extracted, to obtain a satisfactory statistical fit, led in many cases to highly indeterminate factors.

5. An additional problem also manifested itself during

the course of the data reanalysis: "An incidental finding ... was the discovery that oblique rotation often produced doublets in the factor patterns, once m was raised to improve the fit. Such doublets, as is well known (Anderson and Rubin, 1956), correspond to unidentifiable factor patterns, in the sense that the communalities between the two variables which load nonzero on the doublet are arbitrary within certain limits. This is clearly an undesirable situation and it appears to arise with greater frequency than might have been suspected once m is raised so as to satisfy statistical standards" (Schönemann and Wang, 1972, p. 87).

These results show that the factor analyst is often caught in a severe dilemma: either having indeterminate factors or unidentifiable patterns, or a model that does not fit the data.

As the 1970s began, the logical problems inherent in the "estimation" of indeterminate factor scores had still not been resolved. For example, Tucker (1971) examined the relationship of four different factor score estimates to external measures not in the test battery but never mentioned factor indeterminacy. He observed, "The theoretical scores are not available; consequently, several systems for estimating the scores have been proposed." But he never explained why the factor scores are not "available." The notion that factor scores cannot be computed (or are "not available") but must be estimated had thus persisted almost 20 years after Kestelman (1952) had shown that factor scores could indeed be computed. Schönemann and Wang (1972, p. 88) questioned this enduring concept:

> What do such statements mean?
>
> They evidently mean hardly anything as long as we are not told in clear and unambiguous terms what is meant by "factor scores" (as distinct from "factor score estimates"). Upon checking, one finds that the exact meaning of this term is a closely guarded secret. There are good reasons for not defining it: if by "factor scores" one means, as one sometimes does by implication, observations on random variables, then "factor scores" cannot be defined uniquely for the simple reason that the underlying random variables, the "fac-

> tors," cannot be defined uniquely. This is quite different from saying that they cannot be "calculated uniquely" (Horst, 1969, pp. 7–8), which is a minor matter by comparison.

McDonald (1972) took up the challenge of Schönemann and Wang to explain the meaning of factor score "estimation." He defended "traditional treatments of factor scores," maintaining that "seemingly cogent criticisms" of these estimation procedures "by Schönemann and Wang are without foundation" (p. 2). McDonald's major argument began with the premise that one of the sets of factor scores fitting the model is the "correct" one. He then showed that "regression estimates" of the "true" factor scores would, in most cases, correlate more highly with the "true" factor scores than an alternate set of factor scores constructed via Equation (27).

This result, which seemed to provide a new justification for the procedure of "estimating" factor scores with linear combinations of the observed variables, received enthusiastic support in some quarters. The article was accepted for publication in *Psychometrika*. Subsequently, a reviewer (Guttman, 1972) discovered algebraic and semantic imprecisions in the statement of the argument. These created at least the appearance of a self-contradiction, and the article was never published in its original form. Nevertheless, it remains historically significant as perhaps the first substantive attempt to justify the current methods of factor score estimation.

In 1974, McDonald published a paper whose "main aim ... is to show that common factors are not subject to indeterminacy to the extent that has been claimed, because the measure of indeterminacy that has been adopted is ill-founded" (p. 203). He argued that the use of the minimum correlation between alternative equivalent factors "is inconsistent with the foundations of the factor model in probability theory" and that "traditional measures," such as the multiple correlation between the tests and the factors, "yield no disturbing conclusions about the model" (p. 203). McDonald's principal argument was that a "contradiction is contained in any attempt to say that different values of ξ can be associated ('at the same time' is understood throughout this discussion) with one subject.... If an individual is at ξ^+, he is not at ξ^-. That is all there is to it." This position stems, apparently, from a misconception about the nature of mathematical variables. A mathematical variable can be defined (Hays, 1973, p. 29) as "a symbol that can be replaced by any one of the elements of a specified set." Hence, when we say that $T_i = 1$, $T_i = 99$ in the Guttman analogy at the beginning of this chapter, we do not imply the contradictory result $99 = 1$. Rather, we simply imply that both 99 and 1 are members of the solution set. We can define the solution set in factor analysis, just as we can in Guttman's example, and we are perfectly free to apply mathematical operations to its elements. Whether we are computing the difference between 1 and 99 or the difference between two alternative sets of factor scores, we are committing no "mathematical contradiction."

McDonald's paper contains a number of other questionable logical and semantic innovations, such as the notion of an "unobservable random variable" and a "uniquely defined but numerically indeterminate" random variable. Guttman (1975) has described several of these, and he has also raised serious questions about the editorial process that culminated in their publication.

Mulaik (1976) characterized McDonald's attack (1974) on Guttman's minimum correlation index (1955) as a "straw man" argument. Mulaik proved that when alternative solutions for a factor are equally likely to be chosen, the squared multiple correlation ρ^2 for predicting the factor from the observed variables is the average correlation ρ_{AB} between independently selected alternative solutions A and B. Mulaik concluded that the choice between ρ^2 and $2\rho^2 - 1$ "matters little if we keep in mind that these two indices measure different aspects of the same situation."

Green (1976) briefly reviewed some of the more recent factor indeterminacy literature. He criticized the "argumentative, long-winded conclusions" of his predecessors and offered to "clarify and explain the main facts." Ironically, he began by stating (incorrectly) that "factor scores cannot be obtained." He then gave equations for obtaining them. After pointing out (as had Spearman in 1933) the analogy between regression and factor score construction, and (as had Thomson in 1934) the crucial distinction between them, Green dismissed McDonald's argument as "unconvincing and unnecessary," since factor score estimates can be defended as simultaneously "estimating the entire infinite set of possible factor score vectors." He concluded that "a good index of factor score determinacy is the squared multiple correlation of the factor with the observed variables."

Unfortunately, Green's analysis left substantive questions unanswered. If as he insisted, each one of the possible sets of factor scores is "equally true" and has "just as much claim to being the 'true' factor scores as any other," then why is there any need to "estimate" factor scores? To return once again to Guttman's analogy, if we believe that 1 and 99 are equally valid "true scores" for individual *i*, does using an "estimate" of 50 solve our indeterminacy problem?

The relationship between factors and "external variables" (those not included in the test battery) is an area of factor analytic theory that generated virtually no interest before 1976. Most factor analysts apparently assumed that they could learn all they needed to know about factors from their relationships with the original tests. However, this view can be erroneous, for several reasons.

First, when one is evaluating factors, there is no compelling reason to restrict one's interest to the original test battery. In fact, such a restriction is deceptively short-sighted, in view of the original goals of the factor model. If a factor analysis is to discover new random variables that explain a broad "behavior domain," then these new variables, once identified, must relate meaningfully not only to the variables in the test battery but also to substantive variables that may not have been included in the test battery.

Second, the consideration of external variables provides a new perspective for evaluating some previous misconceptions about factor indeterminacy, such as the beliefs that factor indeterminacy is not a problem at the variance-covariance level or that it is only a problem if one wants to obtain factor scores.

Third, the relationship between factors and external variables is crucial to the use of factor analysis as a vehicle for making decisions about people. If factors are to be used as variables in a linear predictive system, then one must ascertain the linear relationship between the factors and the criterion variable.

In exploring some special cases of the factor external variable relationship (Schönemann and Steiger, 1976a), we used a partition of the vectorspace of all deviation score vectors to deduce two theorems. The first theorem says that the factor model implies the existence of external variables which, though perfectly predictable (in the multiple regression sense) from the test battery, are completely unpredictable from any of the possible sets of common factor scores. On the other hand, the same factor model also showed the existence of external variables which, though entirely uncorrelated with the observed scores, are positively correlated with suitably chosen common factor scores.

The second theorem states an even more unexpected result. The common and unique factors of the factor model can always be constructed so as to predict *any* given criterion perfectly, including all those that are entirely uncorrelated with the observed variables. These surprising results further dramatize the arbitrariness of the increment by which factors "go beyond the test space."

Steiger (1976) developed a general theory of the relationship between indeterminate common factors and external variables. He showed that the correlation between a common factor and an external variable is not unique—it can only be specified within a given range. He also gave methods for calculating the upper and lower bounds on these external correlations, as well as the specific factor score vectors associated with these extreme values. In addition, the reanalysis of a number of factor analytic studies showed that external correlations, in practical applications, may have a broad range of unidentifiability. Steiger concluded that factor indeterminacy does have implications at the variance-covariance level, provided one's theoretical perspective is not unnecessarily restrictive.

The 1970s have witnessed a rebirth of interest in the common factor model and its full implications. Many of these implications were available in the past but simply ignored. Other results are new and have often been surprisingly counterintuitive. The indications at present are that research in the area will continue to expand, along with the public awareness of the perplexing problems raised by factor indeterminacy.

SUMMARY AND CONCLUSION

If a single striking fact dominates the history of factor indeterminacy, it is the tendency of the psychometric community to ignore the problem and its implications. Typical instances are (1) the abandonment of the issue during the 1940s and (2) the treatment of the topic of factor score "estimation."

The Thurstonian Era. There is a major gap in the factor indeterminacy literature from 1939 to 1951. These were the peak years for factor analytic research—as Harman (1969) noted, more than 100 articles on factor analysis appeared in *Psychometrika* alone during this period. Yet not one dealt with factor indeterminacy. How could this have happened?

Spearman and Thurstone were clearly the most influential figures in the early history of factor analysis, and their approach to the indeterminacy issue undoubtedly influenced others. Spearman apparently believed the issue had been solved by the regression analogy and limit argument in his 1933 paper. His final work (Spearman and Jones, 1950), published after his death in 1945, contained a forceful restatement of these positions. In contrast, Thurstone, who succeeded Spearman in the 1940s as the dominant figure in factor analysis, never addressed the indeterminacy problem at all. Although this suggests that he was simply unaware of it, a number of facts seem to rule out this explanation. First, Thurstone met with Wilson, Spearman, Holzinger, and others for several days during July 1933, when the indeterminacy issue was attracting maximum attention in the literature. It seems quite likely that the topic was explored during this meeting. Second, Thurstone's two texts (1935, 1947) demonstrated a keen awareness of Wilson's work.

Thurstone frequently emphasized that he considered the assessment of individuals' factor scores of minor importance, compared to the greater goal of discovering and understanding the factors themselves: "The principal purpose is to discover the parameters of factors and something about the nature of the individual differences that they produce. The individual subjects are examined, not for the purpose of learning something about them individually, but rather for the purpose of discovering the underlying factors" (Thurstone, 1947, p. 325). Although in his somewhat cursory discussions of "factor scores" (actually "estimates") Thurstone never directly confronted the indeterminacy issue, on occasion he seemed to allude to it: "Ultimately, we want to be able to appraise each individual as to each of the factors, but this problem raises certain other questions about the domain" (1947, p. 325). "The development of the factorial methods has been a continuous process of compromising between the theoretically complete and ideal solutions, and those solutions that can be made available with a reasonable amount of labor, time, and cost" (1937, p. 511). "The solution to this problem is given . . . for the regression coefficients, but in practical application it is rarely feasible or even desirable to use the theoretically complete solution" (1947, p. 514).

Thurstone had little use for factor scores. Those who do, it would seem, owe their readers a discussion of the facts of indeterminacy and an interpretation of their significance. Remarkably, a number of writers, even as recent a one as Tucker (1971), have managed to discuss factor scores and their "estimation" without fulfilling either obligation.

The Estimation of Factor Scores. Piaggio, having developed the construction equations, knew quite well that factor scores could be computed. He suggested using the determinate regression estimates in place of indeterminate factor scores. Subsequent writers paid more attention to the estimates and less attention to the scores (and their indeterminacy), while attempting to legitimize the "estimation" procedure via the cliché that "factor scores cannot be computed, they can only be estimated."

This notion should have been laid permanently to rest by the work of Kestelman (1952) and Guttman (1955). Instead, it remained alive and popular all through the 1960s. As a result, psychometricians of the era stressed the "estimation" of parameters (factor scores) that could have been readily computed. When Harris (1967) justified such "estimation" with the argument that the factor scores themselves are not unique, and thus of no "practical use," he did not explain why one would even want to estimate a useless, indeterminate parameter.

If such a treatment of factor score "estimation" seemed illogical, no one seemed to mind. And yet the incongruities might have been avoided if writers had given more attention to the significant properties of the parameters they were estimating. Schönemann and Wang noted this, suggesting that factor score estimators pay closer attention to what "exactly it is that is being estimated." Though by no means advocating the procedure, they observed that one could indeed compute factor scores if one desired them. The question "Why estimate (computable, indeterminate) factor scores?" has not yet been satisfactorily answered. Meanwhile, both empirical and theoretical considerations have prompted a more significant question: "Why do a factor analysis at all?

The Factor Model—Pro and Con. Schönemann and Wang's empirical results pointed out the dilemma which frequently confronts the factor analyst: he is usually going to be faced with either a poor statistical fit or indeterminate factors. Considering these problems, he might be better off in the long run with a computationally simpler method, such as component analysis, which "dispenses with all the mathematical and semantical problems which accompany the built-in indeterminancies of the factor model" (1972, p. 88).

Component analysis expresses its latent variables as determinate linear combinations of the observed variables. One writes

$$ {}_p\mathbf{Y}_N = {}_p\mathbf{A}_m\mathbf{X}_N + {}_p\mathbf{E}_N \tag{34} $$

where

$$\mathbf{X} = \mathbf{B}'\mathbf{Y} \tag{35}$$

for some **B** of full-column rank. We have recently examined the properties of "regression component" decomposition, in which **A** in (35) is the regression pattern for predicting **Y** from **X** (Schönemann and Steiger, 1976b). In comparing regression component analysis with factor analysis, we concluded that the former has "a broader range of applicability, greater ease and simplicity of computation, and a more logical and straightforward theory" (p. 175).

Proponents of the factor model have offered a number of reasons for preferring it to component analysis. One is that factor analysis "goes beyond" the original test space, whereas component analysis does not. Although this is correct as far as it goes, it is not entirely clear why factor analysis should yield any additional meaningful information simply because it "goes beyond" the test space. Equation (18) shows that the factors can be expressed as the sum of two terms, one of which is a determinate linear combination of the original variables, the other largely arbitrary. It seems unlikely that this arbitrary term adds anything of value to the factors. It could be that the regression "estimates" contain meaningful information not because they "estimate" the factors but because, like components, they are linear combinations of the original tests. If these regression "estimates" turn out in practice to be very similar to the components, then it would seem that the factor model exceeds the component method by little more than "an increment of error."

Velicer (1972) has empirically compared factor analysis with principal components analysis and rescaled image analysis (two special cases of component analysis). He found very little difference in the results given by the three methods, in either the patterns or the scores. The component scores matched the typical factor score estimates very closely. These results seem to support the observations of a number of authors, such as Morrison (1967), who in practical applications have found that the methods tend to give very similar results.

Clearly, numerical examples can be constructed in which one of the methods, factor analysis or component analysis, yields a correct answer while the other method does not. Such examples merely capitalize on the theoretical dichotomy between components (which must be expressible solely as linear combinations of the observed variables) and factors (which can never be so expressed) and hence prove nothing about the relative merits of the two methods. [It is interesting to note that Wilson and Worcester (1939) gave such an example as a demonstration of the limits of component analysis relative to factor analysis.] The more significant issue is whether, in the situations *where they are commonly used,* either method offers significant advantages (relative to its theoretical defects) over the other.

Principal components analysis has often been characterized as a convenient but inferior "approximation" of factor analysis. Instead, it might be more realistic to view factor analysis as a computationally difficult, theoretically problematic, approximation of component analysis.

Factor Indeterminacy and the Conduct of Science. The practical consequences of factor indeterminacy for the modern user are minor, compared with the negative impact the problem has had on the field of psychometrics. [Indeed, Wilson and Worcester (1939) argued that factor analysis could continue to provide some useful information in the face of indeterminacy.] The handling of the issue provides a graphic illustration of how science can function suboptimally.

Almost 50 years have passed since Wilson first introduced the indeterminacy issue. It is still the focus of debate. And yet thousands who were encouraged to use factor analysis were never told about factor indeterminacy. Men of considerable talent spent countless hours refining the technical aspects of factor analysis, while the indeterminacy problem remained obscured.

A science can progress only if its practitioners are willing to confront crucial and difficult theoretical issues head on, rather than postpone them for some future generation. Such theoretical groundwork is often tortuous, but it forms the necessary foundation for any rational science.

Factor indeterminacy poses some difficult problems for those who support the factor model as the superior rationale for the scientific analysis of data. However difficult, these problems should be publicized and debated now. It would indeed be ironic if, given the current level of knowledge, the factor indeterminacy issue were allowed to fade again into obscurity.

REFERENCES

ANDERSON, T. W., and RUBIN, H. "Statistical Inference in Factor Analysis." In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press, 1956.

CAMP, B. H. "The Converse of Spearman's Two-Factor Theorem." *Biometrika,* 1932, *24,* 418–427.

CAMP, B. H. "Spearman's General Factor Again." *Biometrika,* 1934, *26,* 260–261.

DODD, S. C. "The Theory of Factors, I." *Psychological Review,* 1928a, *35,* 211–234.

DODD, S. C. "The Theory of Factors, II." *Psychological Review,* 1928b, *35,* 261–279.

GARNETT, J. C. M. "General Ability, Cleverness, and Purpose." *British Journal of Psychology,* 1919, *9,* 345–366.

GARNETT, J. C. M. "On Certain Independent Factors in Mental Measurements." *Proceedings of the Royal Society A,* 1919, *96,* 91–111.

GARNETT, J. C. M. "The Single General Factor in Dissimilar Mental Measurements." *British Journal of Psychology,* 1920, *10,* 242–258.

GORSUCH, R. L. *Factor Analysis.* Philadelphia: Saunders, 1974.

GREEN, B. F. "On the Factor Score Controversy." *Psychometrika,* 1976, *41,* 263–266.

GUTTMAN, L. "The Determinacy of Factor Score Matrices With Implications for Five Other Basic Problems of Common-Factor Theory." *British Journal of Statistical Psychology,* 1955, *8,* 65–81.

GUTTMAN, L. "'Best Possible' Systematic Estimates of Communalities." *Psychometrika,* 1956, *21,* 273–285.

GUTTMAN, L. Personal communication, 1972.

GUTTMAN, L. "The Green-McDonald Proof of the Nonexistence of Factor Analysis." Unpublished manuscript, 1975.

HARMAN, H. H. *Modern Factor Analysis.* Chicago: University of Chicago Press, 1960.

HARMAN, H. H. *Modern Factor Analysis.* (2nd rev. ed.) Chicago: University of Chicago Press, 1967.

HARMAN, H. H. "Retrospective Anticipation." *Psychometrika,* 1969, *34,* 407–420.

HARRIS, C. W. "On Factors and Factor Scores." *Psychometrika,* 1967, *32,* 363–379.

HART, B., and SPEARMAN, C. "General Ability: Its Existence and Nature." *British Journal of Psychology,* 1912, *5,* 51–84.

HAYS, W. L. *Statistics for the Social Sciences.* New York: Holt, Rinehart and Winston, 1973.

HEERMANN, E. F. "The Geometry of Factorial Indeterminacy." *Psychometrika,* 1964, *29,* 371–381.

HEERMANN, E. F. "The Algebra of Factorial Indeterminacy." *Psychometrika,* 1966, *31,* 539–543.

HEYWOOD, H. B. "On Finite Sequences of Real Numbers." *Proceedings of the Royal Society of London,* 1931, *134,* 486–501.

HOLZINGER, K. J., and HARMAN, H. H. *Factor Analysis.* Chicago: University of Chicago Press, 1941.

HOWE, W. G. "Some Contributions to Factor Analysis." Report No. ONRL-1919. Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1955.

IRWIN, J. O. "On the Uniqueness of the Factor *g* for General Intelligence." *British Journal of Psychology,* 1932, *22,* 359–363.

IRWIN, J. O. "On the Indeterminacy in the Estimate of *g*." *British Journal of Psychology,* 1935, *25,* 393–394.

JÖRESKOG, K. G. "Some Contributions to Maximum Likelihood Factor Analysis." *Psychometrika,* 1967, *32,* 443–482.

KESTELMAN, H. "The Fundamental Equation of Factor Analysis," *British Journal of Psychology,* 1952, *5,* 1–6.

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.

LEDERMANN, W. "The Orthogonal Transformation of a Factorial Matrix into Itself." *Psychometrika,* 1938, *3,* 181–187.

MCDONALD, R. P. "Some Uncertainties about Factor Indeterminacy." Unpublished manuscript, 1972.

MCDONALD, R. P. "The Measurement of Factor Indeterminacy." *Psychometrika,* 1974, *39,* 203–221.

MCDONALD, R. P., and BURR, E. J. "A Comparison of Four Methods of Constructing Factor Scores." *Psychometrika,* 1967, *32,* 381–401.

MCNEMAR, Q. "The Factors in Factoring Behavior." *Psychometrika,* 1951, *16,* 353–359.

MEYER, E. P. "On the Relationship Between Ratio of Number of Variables to Number of Factors and Factorial Determinacy." *Psychometrika,* 1973, *38,* 375–378.

MORRISON, D. F. *Multivariate Statistical Methods.* New York: McGraw-Hill, 1967.

MULAIK, S. A. *The Foundations of Factor Analysis.* New York: McGraw-Hill, 1972.

MULAIK, S. A. "Comments on 'The Measurement of Factorial Indeterminacy.'" *Psychometrika,* 1976, *41,* 249–262.

PIAGGIO, H. T. H. "The General Factor in Spearman's Theory of Intelligence." *Nature,* 1931, *127,* 56–57.

PIAGGIO, H. T. H. "Three Sets of Conditions Necessary for the Existence of a *g* That Is Real and Unique Except in Sign." *British Journal of Psychology,* 1933, *24,* 88–105.

PIAGGIO, H. T. H. "Approximate General and Specific Factors Without Indeterminate Parts." *British Journal of Psychology,* 1935, *25,* 485–489.

PIAGGIO, H. T. H., and DALLAS, A. E. M. M. "An Analysis of Recent Tests of the Two-Factor Theory." *British Journal of Psychology,* 1934, *25,* 217–220.

SCHÖNEMANN, P. H. "The Minimum Average Correlation Between Equivalent Sets of Uncorrelated Factors." *Psychometrika,* 1971, *36,* 21–30.

SCHÖNEMANN, P. H. "Erratum." *Psychometrika,* 1973, *38,* 149.

SCHÖNEMANN, P. H., and STEIGER, J. H. "A Note on Factor Score Validity." Paper presented at the Symposium on Application in Statistics, Dayton, Ohio, June, 1976a.

SCHÖNEMANN, P. H., and STEIGER, J. H. "Regression Component Analysis." *British Journal of Mathematical and Statistical Psychology,* 1976b, *29,* 175–189.

SCHÖNEMANN, P. H., and WANG, M. M. "Some New Results on Factor Indeterminacy." *Psychometrika,* 1972, *37,* 61–91.

SPEARMAN, C. "General Intelligence, Objectively Determined and Measured." *American Journal of Psychology,* 1904, *15,* 201–293.

SPEARMAN, C. "Correlation Between Arrays in a Table of Correlations." *Proceedings of the Royal Society A,* 1922, *101,* 94–100.

SPEARMAN, C. *The Abilities of Man.* New York: Macmillan, 1927.

SPEARMAN, C. "The Uniqueness of *g*." *Journal of Educational Psychology,* 1929, *20,* 212–216.

SPEARMAN, C. "The Uniqueness of *g*." *Journal of Educational Psychology*, 1933, *24*, 106–108.

SPEARMAN, C. "The Factor Theory and Its Troubles, IV: Uniqueness of *g*." *Journal of Educational Psychology*, 1934, *25*, 142–153.

SPEARMAN, C., and JONES, L. W. *Human Ability: A Continuation of "The Abilities of Man."* London: Macmillan, 1950.

STEIGER, J. H. "The Relationship Between External Variables and Indeterminate Factors." Unpublished doctoral dissertation, Purdue University, 1976.

THOMSON, G. H. "A Hierarchy Without a General Factor." *British Journal of Psychology*, 1916, *8*, 271–281.

THOMSON, G. H. "The Hierarchy of Abilities." *British Journal of Psychology*, 1919, *9*, 337–344.

THOMSON, G. H. "The Meaning of *i* in the Estimate of *g*." *British Journal of Psychology*, 1934, *25*, 92–99.

THOMSON, G. H. "The Definition and Measurement of *g* (General Intelligence)." *Journal of Educational Psychology*, 1935, *26*, 241–262.

THOMSON, G. H. *Factorial Analysis of Human Ability*. London: University of London Press, 1948.

THURSTONE, L. L. *The Vectors of Mind*. Chicago: University of Chicago Press, 1935.

THURSTONE, L. L. *Multiple Factor Analysis*. Chicago: University of Chicago Press, 1947.

TUCKER, L. "Relations of Factor Score Estimates to Their Use." *Psychometrika*, 1971, *36*, 427–436.

VELICER, W. "An Empirical Comparison of Factor Analysis, Image Analysis, and Principal Component Analysis." Unpublished doctoral dissertation, Purdue University, 1972.

WILSON, E. B. "Review of 'The Abilities of Man, Their Nature and Measurement' by C. Spearman." *Science*, 1928a, *67*, 244–248.

WILSON, E. B. "On Hierarchical Correlation Systems." *Proceedings of the National Academy of Sciences*, 1928b, *14*, 283–291.

WILSON, E. B. "Review of 'Crossroads in the Mind of Man: A Study of Differentiable Mental Abilities' by T. L. Kelley." *Journal of General Psychology*, 1929a, *2*, 153–169.

WILSON, E. B. "Comments on Professor Spearman's Note." *Journal of Educational Psychology*, 1929b, *20*, 217–223.

WILSON, E. B., and WORCESTER, J. "A Note on Factor Analysis." *Psychometrika*, 1939, *4*, 133–148.

WOLFLE, D. *Factor Analysis to 1940*. Psychometric Monographs, No. 3. Chicago: University of Chicago Press, 1940.