

Four Queries About Factor Reality

Paul E. Meehl

University of Minnesota

When I first learned about factor analysis (as an undergraduate over a half-century ago) there were four methodological questions that troubled me. They still do. Although I have used factor analysis in research, with excellent results, it is disconcerting to realize that these basic issues remain unsettled, and are rarely examined with the incisiveness and depth they deserve. I have neither the expertise nor temerity to “answer” them, so I shall only try to put them clearly and forcibly. They are *not*, I am sure, answerable by fiddling with the mathematics, which may be why mathematical statisticians do not generally regard factor analysis highly.

1. **What sort of existential status (Feigl, 1950) do we—and ought we—impute to factors?** For a scientific realist, a factor is presumably a physical entity possessing a quantitative property. (I exclude metaphysical dualism, held by hardly any psychologists.) The physical entity exists in *the person*, hence, in the brain. (Where else? Cognitive, affective, and motivational states, events, and dispositions are not located in the kidney!) The general intelligence factor *g* is “in” the CNS. The heritable component of *g* could perhaps be identified with “number of polygenes of the (+) bright kind.” Is such an identification objectionable?

Some think that the elements (components, summands, “units”) of a real physical factor must be qualitatively homogeneous. Must they? A properly done factor analysis of battles would surely reveal *information* as a potent factor, sometimes critical (cf. Col. Hentsch’s unclear message to Gen. von Kluck, which led to the first Battle of the Marne and probably lost World War I for Germany). Information can be expressed by a number (e.g., bits $\times$ distance $\div$ time), and it doesn’t matter *how* it is transmitted. But the physical modes are qualitatively unlike (runner, motorcyclist, carrier pigeon, field telegraph, telephone, radio, heliograph). Suppose the heritable component of *g* involves only 5 additive chromosomal loci, controlling such microstructural properties as number of neurons, density of dendritic twigs, number and length of collaterals, spatial distribution of terminal buttons over the synaptic scale, average number of cells in a module, etc. If independent, the additivity gives us 10 events, and if $p \simeq q$ at each locus, expansion of $(p + q)^n$ yields a frequency polygon practically indistinguishable from the Gaussian function. Further, if 10 variables have positive manifold, how they are *weighted* in a composite hardly matters. It seems that a mathematical factor can correspond to a causally efficacious composite whose elements are qualitatively unlike. Is this objectionable for some reason?

2. **What is the purpose of factor analysis if one is a fictionist rather than a realist?** Fictionists sometimes say that factor analysis is merely a convenient, economical way of representing a batch of correlations, and makes no stronger claim for the status of factors. This strikes me as odd, even a bit disingenuous. I have a matrix of correlation coefficients between tests. I then present (instead? along with?) a

factor matrix, informing you that if you multiply the factor matrix by its transpose you will (almost) recapture the correlations. Where is the economy? Seems like a lot of work to generate some (non-denoting!) numbers which, if manipulated in a rather complicated way (matrix algebra), does not quite yield the original denoting numbers. And why would a consistent fictionist even be concerned with the rotation problem, when all positions of the reference axes are merely mathematical transformations, representing the same empirical facts in different ways (coordinates)?

The instrumentalist may reply “We use the factors to help us build better tests.” If the “insight” into what underlying psychisms enter into test items of such-and-such kinds is spurious—if there are in reality no such mental factors as those identified—what reason is there to expect new items built in reliance on the factor analysis to “do better”? And just what is it they are supposed to “do better *at*,” on the fictionist view? What evidence shows that the test-builder gets superior results from factor analysis to what would be attained by simply contemplating the pattern of correlations? (“What does introspection suggest as the mental process shared by tests X and Y, close together on the hypersphere, and somewhat less by Z, a bit more distant from both of them?”)

Robert C. Tryon once told me that he had cluster analyzed many data sets which had been factor analyzed by Thurstone and others. He claimed that his cluster analysis method (simpler, easier, and no rotation problem) usually gave results very close to Thurstone’s, and when the methods differed markedly, Tryon’s clusters were easier and clearer to interpret (“understand psychologically”) than Thurstone’s factors.

I take it as noncontroversial that in a purely instrumental task, predicting an “operationally specified” criterion of intrinsic pragmatic interest, there is no reason to employ factor analysis. Instead one uses an appropriate straightforward optimizing procedure (multiple regression, discriminant function, actuarial table, etc.). There is no point in path analysis either, unless one expects to conclude with causal imputation or denial.

3. **How dependent are results upon the choice of tests?** Everyone agrees with the clichés “You cannot get more than what you put in,” “Garbage in, garbage out,” “Test building or selection is more important than the formalism.” I am not challenging these accepted principles. But I do not find them helpful in answering the question, “*How dependent?*” A simple point first: All solutions to the rotation problem utilize functions of the distribution of factor loadings, which are in turn functions of the number and density of tests on regions of the hypersphere. Everyone knows that if a psychological factor is not “represented” by even a single test in the battery, it obviously cannot emerge from the analysis. But isn’t the opposite effect equally important? If one is clever and indefatigable, one can modify tests (materials, instructions, scoring, metric, transformations) in a sparse region of the hypersphere so as to get more representation of a tentative factor, and by continuing that process can split the factor into new factors. Is there a “rational” termination to this densifying process? I am aware of none. One nonarbitrary stopping rule would be to quit concocting new tests when the correlations approach being equivalent forms reliabilities. Nobody has ever tried that, presumably because it would mean concocting not hundreds but *thousands* of tests, even in a fairly limited domain (e.g., cognition, psychopathology). Are there good reasons for expecting that such a process would (always, usually, or *ever*) result in such clear differences in *achievable* densification, defined by qualitative and quantitative

closeness in the phenotypic space, as to convince rational minds that the factors were really there, “waiting to be found” if we densify nonarbitrarily? Perhaps Guilford’s 120 factors of intellect are a feeble approximation of this. Have psychologists been persuaded by him? Had he persisted in test building, setting aside his a priori categories, could he have found 240 or 480 factors? No one knows. My conjecture is that he could have. This is surely different from the chemist’s situation, where—once the post-alchemists hit upon the right phenotypic properties to attend to (weight, density, conductivity, boiling point, combining powers, etc.) they were sure to identify 102 elements rather than 4 or 400.

4. **What is the rôle of parsimony?** Does a particular solution of the rotation problem (e.g., maximizing the dispersion of loadings as a quantitative generalization of Thurstone’s simple structure rule to maximize quasi-zero loadings) rely on parsimony? Logicians cannot provide a definition of ‘simplicity,’ nor a guarantee that the world is simple. The only two variants of parsimony I could defend are Popper’s and Occam’s. In general, simpler hypotheses are more falsifiable, hence preferable to Popperians. Occam’s *Entia non sunt multiplicanda praeter necessitatem* (I’m told he never said it thus) can be taken to mean “Do not bother inventing explanations for something that is already explained.” Do either of these variants of parsimony lead directly to a factor analytic rotation criterion? I do not know, but I have not seen it done, and I rather doubt it is possible.

Reference:

Feigl, H. (1950). Existential hypotheses: Realistic versus phenomenistic interpretations. *Philosophy of Science*, 17, 35–62. Reprinted in his *Inquiries and provocations: Selected writings 1929–1974* (pp. 192–223) (R. Cohen, Ed.). Boston: D. Reidel, 1981.