CPC, 1997, 16 (6), 740-749

COMMENTARY ON « FAMOUS ARTEFACTS »
(P. H. SCHÖNEMANN)

Exactly what *is* an artefact?

Michael D. Maraun

Simon Fraser University

Professor Schönemann's paper, «Famous artefacts: Spearman's hypothesis», is a brave and fascinating analysis of a contentious issue. The current article is his most recent in a series of mathematical analyses centering on the work of Arthur Jensen, and ranging from the implications for his claims of factor analytic indeterminacy (Schönemann, 1987, 1990), to the meaning of the so-called Spearman correlations (Schönemann, 1985, 1992). I use the term brave because very few seem willing to critically evaluate Jensen's work, especially when it comes to technical issues. I agree with the majority of the conclusions from Schönemann's past work, including his observation that far too many of Jensen's claims have attached to them a convenient semantic imprecision (e.g., the calling of the first principal component, *g*). But despite the quality of Schönemann's work, I do not believe that 'Famous artefacts' constitutes a fair analysis of Spearman correlations.

Correspondence should be sent to Dr Michael Maraun, Department of Psychology, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 (e-mail: Maraun@sfu.ca).

What is an artefact?

> « For instance, the Roman astronomers had to be convinced that the things that they saw through the telescope were not optical illusions produced by the instrument....Furthermore, the whole question of the relation of sensory experience, aided by instruments, to 'reality' is by no means simple, and Galileo meditated deeply on the problem » (Segre, 1969, p.20).

What is an artefactual result? The answer is by no means straightforward. Schönemann (Note 1) accepts the following definition: « Structure or phenomenon of artificial or accidental origin encountered during an observation or experiment bearing on a natural phenomenon » (Nouveau Petit Larousse, 1972, p. 69). In Webster's New College Dictionary (1990) one is informed that an artefact is « A product (as a structure on a prepared microscope slide) of artificial character due to extraneous (as human) agency. » Artefactual results seem to be describable according to the following kind of scenario. One studies an object/phenomenon (O), with an instrument (I) that produces a result (R). If R is a representation of properties of O then it provides information about O. One can rightly, in certain contexts, claim to have made an empirical discovery on the basis of R. One may hypothesize something about O, and use R as evidence for or against the hypothesis. If, on the other hand, R is not a representation of properties of O, but is a byproduct of the use of the instrument itself, then it is an artefact (A). Unfortunately, the relation between the object/phenomenon of interest and the instrument by which it is represented is by no means simple.

How does one decide whether R is artefact or representation? Clearly, there exist many different domain specific procedures that may be employed to arrive at a judgment. Depending on the context, one might change the settings of the instrument, or temporarily employ a different instrument. Sometimes the inappropriate choice or operation of an instrument will lead to the generation of an artefactual result. If R remains after such a change, then it might provisionally be taken as a representation of O. One might examine different objects under the same instrument set-up (e.g., as with a microscope lens), and if R is found regardless of which object is studied, then one might suspect it to be an artefact.

The question is whether Spearman correlations may rightly be called artefactual results. It seems to me that it is rather difficult to translate the terms of the above scenario into the data analytic terms that are the basis for the study of Spearman correlations. Nevertheless, in such a data analytic context, the place of instrument might be seen as occupied by an *analytic framework* of 'data theory'. The analytic framework contains the building blocks in terms of which the problem is defined. It includes details regarding the populations under study, the specifics of data collection (e.g., the sampling scheme involved), the correspondence relations between phenomena of interest and data analytic concepts, and the data analytic models employed to represent the data. It fixes what is meant by an *outcome* of a study, and provides a vocabulary for the phrasing of hypotheses and conclusions.

Consider the case of difficulty factors in the context of linear factor analysis (McDonald & Ahlawat, 1974). Here, one is interested in 'observing' the relationships among a set of dichotomous items (O). However, the object under study is located in N-dimensional euclidean space (N = number of subjects), and so cannot be observed unaided. One requires an instrument to generate a low-dimensional projection (representation). One could choose linear factor analysis for this purpose. The instrument would then be the usual embedding framework of statistical concepts one comes across in a consideration of multivariate data, and those particular to factor analysis. Now, the factor analytic result is an r dimensional projection, in which r is much smaller than N. But now assume that, as part of this result, one finds a factor that is highly correlated with the mean vector of the items (R). One might then declare R to be an artefact because it has not to do with O. How does one know this? Because, generally speaking, the degree to which R manifests itself is related to the value of the mean vector (Olsson, 1979), and the mean vector is univariate information (and so should have no bearing on O, *relationships* among the items). Furthermore, the use of a more theoretically appropriate instrument (non-linear factor analysis) rids one of the artefactual result (McDonald & Ahlawat, 1974). Consider a second example, the estimation of the parameters in a statistical model. One would like to take the estimates as representing features of the phenomenon under study. But how does one know that the estimates can be taken in this way, and are not simply a function of the starting values chosen? The obvious answer is that one should change the starting values a number of times and observe whether the same estimates are obtained. If radically different estimates are obtained with different starting values, one should take the estimates to be artefactual. This example may, in fact, be closer to a data analytic sense of artefactuality than the previous one.

I don't think that Schönemann wishes to claim that the Spearman correlations are artefactual. In my opinion, he has something different in mind which is closer to *trivial dependency*, than artefactuality. This sense is in keeping with the following:

> « The importance that Jensen (1985) attaches to this hypothesis is attested to by his statement: "[the] hypothesis that the magnitudes of black-white mean differences on various mental tests are directly related to the test *g* loadings, if fully substantiated, would be an important and unifying discovery" (p. 197) ... I shall substantiate this proposition – but not its importance – by showing it to be but a mathematical consequence of the *g* hypothesis and not at all in need of empirical evidence » (Guttman, 1992, p. 196).

The idea is that, given the analytic framework, the Spearman correlation may be a necessary consequence of other empirical results, and not itself an independent empirical discovery. It may not be the case that it can be independently *hypothesized* because it may just be a necessary consequence of a sub-set of other empirical results *plus* aspects of the construction of the problem. When these other results are obtained, the Spearman correlation hypothesis may not be falsifiable. This seems to me to be a very different thing than artefactuality, artefactuality suggesting that, in some way, a result does not truly represent features of the object under study. In a data analytic problem, one might define this sense of *trivial dependence* as follows:

> ***Definition***: Given an appropriate analytic framework, a result A is trivially dependent on result B if A is a necessary consequence of B.

It is not hard to generate examples of this kind of dependency. For the case of two dichotomous items, one might think of the relationship between the univariate marginals and the range over which ϕ may vary.

Whether, in a given context, the existence of such a dependency is damaging depends upon the roles that A and B were to play. For example, if A is dependent on B, then, given B, it would be trivial to hypothesize A. Given B, the hypothesis of A would not be falsifiable. Similarly, given B, it would be silly to put forth A as a major, independent empirical finding. For the case of two dichotomous items, it would be much like observing $P_1=.7$ and $P_2=.3$ and then hypothesizing that $\phi<.43$. To paint Spearman correlations in this way was Guttman's aim in his 1992 paper (just take A to be 'Spearman correlation' and B to be 'Single common factor model holds'). Regardless, there are certain cases in which dependency may lead to even greater trouble. For example, if one held result D as a possible explanatory prop for result C, and it turned out that D was dependent on C, then one would be in a bind. This is the possibility that is most menacing with regard to Spearman correlations.

The analytic framework of the Spearman correlation issue

Regardless of whether one accepts this distinction between artefactuality and dependency, one must specify the analytic framework that underlies the study of Spearman correlations. The Spearman correlation issue features a P-vector of random variates, $\underline{X}$, each variate representing an intelligence test, this vector distributed on R^P. There exist two distinct populations, whites and blacks, occurring with proportions π_1 and π_2. Also featured are three covariance matrices, three mean vectors, and the eigen-structures of the covariance matrices. The unconditional density of $\underline{X}$ is therefore a mixture:

$$1)\ f(\underline{X}) = \pi_1 * f(\underline{X}|1) + \pi_2 * f(\underline{X}|2).$$

with $f(\underline{X}|1)$ and $f(\underline{X}|2)$ the conditional densities of populations 1 and 2, respectively. All first and second moments are assumed to exist, and then

$$2)\ \underline{\mu} = \pi_1 * \pi_1 + \pi_2 * \pi_2 \text{ and}$$
$$3)\ \Sigma = \pi_1 * \Sigma_1 + \pi_2 * \Sigma_2 + \pi_1\pi_2 * \underline{dd}',$$

in which $\underline{\mu}_1$ and $\underline{\mu}_2$ are the mean vectors for populations 1 and 2, Σ_1 and Σ_2 are the covariance matrices for populations 1 and 2, and $d = (\underline{\mu}_1 - \underline{\mu}_2)$. The analytic framework for this problem includes no specification as to the specific form of the unconditional or conditional distributions of $\underline{X}$. Their specific forms is an empirical question. Even if, in practice, $f(\underline{X}|1)$ and $f(\underline{X}|2)$ turn out to be multinormal, $f(\underline{X})$ will not in general be so. Let this analytic framework be called B1.

Given B1, the hypotheses are made that

4) $\cos^2_{\theta(\underline{v}1,\underline{d})} >> 0$ in which $\Sigma\underline{v}_1 = \lambda_1\underline{v}_1$, $\underline{v}_1'\underline{v}_1 = 1$
(Spearman level 1), and
5) $\cos^2_{\theta(\underline{w}1,\underline{d})} >> 0$, $\cos^2_{\theta(\underline{z}1,\underline{d})} >> 0$
in which $\Sigma_1\underline{w}_1 = \delta_1\underline{w}_1$, $\underline{w}_1'\underline{w}_1 = 1$, $\Sigma_2\underline{z}_1 = \phi_1\underline{z}_1$, and $\underline{z}_1'\underline{z}_1 = 1$,
(Spearman level 2).

Schönemann's claim

Schönemann claims that both the level 1 and level 2 correlations are artefacts. Once again, I suggest that his claim should be paraphrased as one about dependencies. Regardless, Schönemann proves that (4) and (5) are a necessary consequence of:

6) $\underline{X} \sim N(\underline{\mu}, \Sigma)$
7) Populations 1 and 2 are defined as the regions on either side of the cutting plane through the centroid, and orthogonal to $\underline{v}_1$, i.e., if $(\underline{X}-\underline{\mu})'\underline{v}_1 > 0$ then $\underline{X} \subset P1$, otherwise $\underline{X} \subset P2$
(secondary consequence) $\Sigma_1 = \Sigma_2$
8) Σ_1 and Σ_2 contain only positive elements.

Has he revealed the existence of a damaging dependency? I do not believe so, for Schönemann has not chosen a reasonable analytic framework. Specifically, (6), (7), and (8) are severe restrictions on B1, in which the conditional densities of (1), $f(\underline{X}|1)$ and $f(\underline{X}|2)$, are truncated normals, the unconditional density is multivariate normal, and the conditional covariance matrices of (3) are equal and "positive manifold". Schönemann's conditions are, in fact, testable *hypotheses* within B1. Therefore, they themselves do not constitute an appropriate framework. I can think of no reason why (6), (7), and (8) should be the starting point for an analysis of the Spearman correlation issue.

Schönemann, on the other hand, doubts

« ... how anyone could question multinormality as a default assumption. Multinormality undergirds virtually all multivariate statistics practiced in the social sciences, including maximum likelihood factor analysis, LISREL, tetrachoric correlations, corrections for attenuation, most likelihood ratio tests, most classical test theory, etc., etc. Why should this ubiquitous and universally accepted premise all of a sudden become suspect when it is invoked to refute counter-intuitive claims of social import? » (Schönemann, 1997, p. 689).

He erects a similar defense of (7) and (8). But this is placing the cart well before the horse. It is true that multinormality is routinely assumed in the formulation and use of statistical models, but this fact is wholly irrelevant to the current issue. The choice of an appropriate analytic framework for the consideration of an empirical issue is in no way constrained by the (often expeditious) considerations that enter into the development and use of statistical models. Statistics, when used properly, is the servant of science, and not the other way around. The choice of an appropriate analytic framework has but one master: The adequate exploration of the empirical issue of interest. If then multinormality *is* required to power a particular statistical procedure, then the procedure is inappropriate for the study of Spearman correlations.

Do there exist damaging dependencies?

Given an appropriate analytic framework, i.e., B1, do there exist dependencies that would be damaging to the status of Spearman correlations? To answer this question, one must remember what purpose the Spearman correlations were to serve. As is clear from Jensen's 1985 article, and Guttman's 1992 commentary, the issue is whether a Spearman correlation, or its squared cosine counterpart, may be used as an explanatory prop for $\underline{d}$, the Black-White mean difference vector. This potential role would be undermined if, e.g., the Spearman correlation was dependent on $\underline{d}$. In my opinion then, there exist difficulties with Spearman level 1, but not with level 2.

1. Spearman level 1. Assume B1. Then as

$$\underline{d} \rightarrow \infty,\ \cos^2{}_{\theta(\underline{v}1,\underline{d})},\ \text{i.e., } \underline{v}_1 \text{ and } \underline{d} \text{ become collinear.}$$

The proof is obvious from the geometry of the 2-component mixture distribution. Remember that

$$\Sigma = (\pi_1{*}\Sigma_1 + \pi_2{*}\Sigma_2) + \pi_1\pi_2{*}\underline{dd}' = A + B$$

Following Schönemann (1985), B is a rank one matrix with eigenvalue $\pi_1\pi_2\underline{d}'\underline{d}$, and eigenvector $\underline{d}$. As $\underline{d}$ gets large, matrix B increasingly dominates Σ, since A is a constant. Hence, as $\underline{d}$ gets large, it will become the first eigenvector of Σ, and so becomes collinear with $\underline{v}_1$. This is damaging because the level 1 result is a finding dependent upon the result it was supposed to explain, i.e., large Black-White mean differences.[1] To put this another way, the mere fact that the populations were very different in their mean vectors, i.e., a large $\underline{d}$ existed, would lead one to conclude that the 'source' of this difference was *g*. Hence, as long as $\underline{d}$ was large enough, the 'hypothesis' that *g* 'explains' these mean differences would not be falsifiable. Schönemann makes this point in both his 1985 and current papers.

Figures 1a, 1b, and 1c illustrate this situation for a bivariate mixture of two populations (densities unspecified). In all three figures, $\pi_1 = .3$, $\pi_2 = .7$, $\underline{\mu}_2 = [2,2]$, and

$$\Sigma_1 = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}; \Sigma_2 = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix}$$

However, in Figure 1a, μ_1 = [2.04 2.16], making $\underline{d}$ = [.04 .16], $\underline{v}_1$ = [-.706 .706], and $\cos^2{}_{\theta(\underline{v}1,\underline{d})} = .261$. In Figure 1b, μ_1 = [2.4 3.6], making $\underline{d}$ = [.4 1.6], $\underline{v}_1$=[-.569 .823], and $\cos^2{}_{\theta(\underline{v}1,\underline{d})} = .436$. Finally, in Figure 1c, μ_1 =[4 10], making $\underline{d}$ = [2 8], $\underline{v}_1$=[.192 .981], and $\cos^2{}_{\theta(\underline{v}1,\underline{d})} = .996$.

2. Spearman level 2. Given B1, does there exist such a dependency for the case of the Spearman 2 correlation. I think that it is evident from the geometry of the 2-component mixture distribution that there does not. Specifically, the orientations of $\underline{d}$, $\underline{w}_1$, and $\underline{z}_1$ in R^P are not restricted in *any* way. Hence, collinearity is most certainly a finding independent of that which it would be taken to explain, i.e., a large $\underline{d}$.

Now, Schönemann's (6), (7), and (8) need not arise empirically within B1. However, what would it mean if, as Schönemann claims (Note 6a), they *did* consistently arise empirically: « Among the former, I count multinormality, which is critical for Level II and, I believe, eminently plausible » (Schönemann, 1997, p. 689). Equivalently, what if, for the populations under study, (1), (2), and (3) turned out empiri-

---

1. One might counter that a large $\underline{d}$ is not *necessary* for Spearman 1, but this does not help since the existence of a large $\underline{d}$ is not in question.

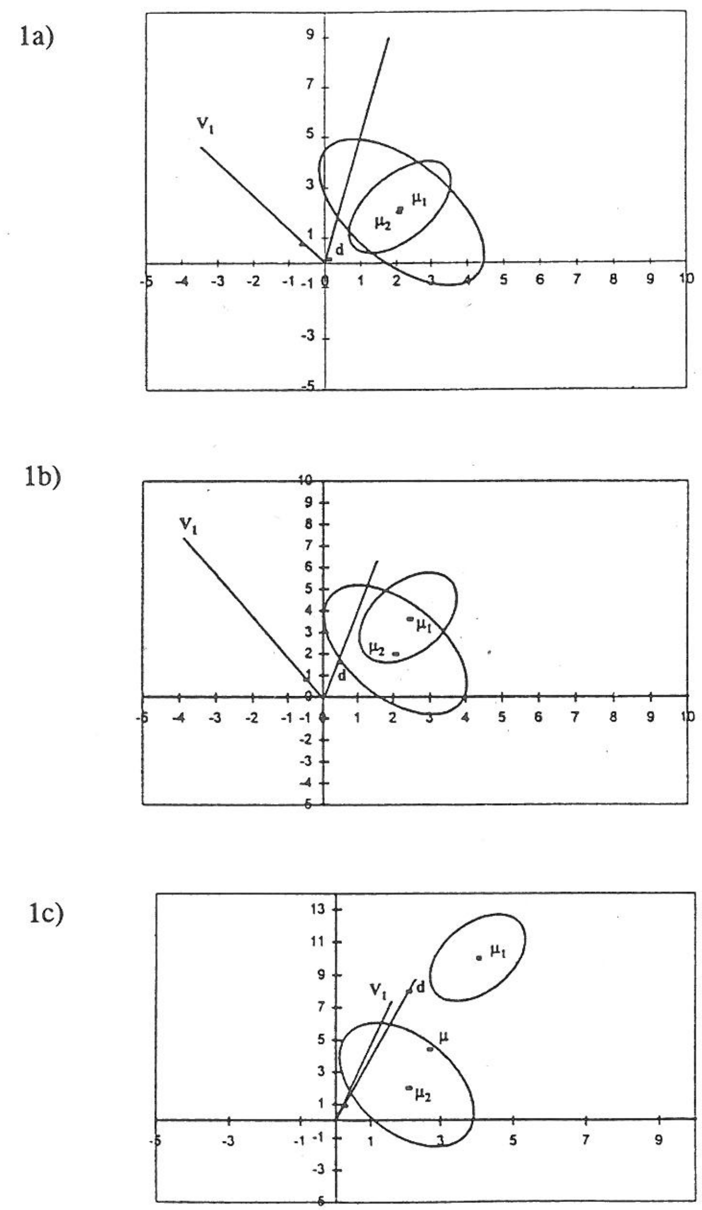

Figure 1. Spearman Level 1 (Cosines) as $\underline{d} = (\mu_1 - \mu_2)$ *increases: Bivariate mixture.*

cally to be (6), (7), and (8). Would *this* then signal the existence of a damaging dependency? Not in my opinion. Certainly, an empirical result may arise for many different reasons, and Schönemann has picked out from a very broad class a sufficient condition for the Spearman 2 correlation. However, what is important is that one may still observe a large d, and then non-trivially hypothesize a Spearman level 2 correlation.

ACKNOWLEDGEMENTS

This research was supported in part by a SSHRC small grant awarded to the author. I would like to express my gratitute to Ray Koopman for helpful comments on an earlier version of this manuscript.

REFERENCES

Guttman, L. (1992). The irrelevance of factor analysis for the study of group differences. *Multivariate Behavioral Research, 27* (2), 175-204.

McDonald, R., & Ahlawat, K. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology, 27,* 82-99.

Schönemann, P. H. (1985). On artificial intelligence. *The Behavioral and Brain Sciences, 8* (2), 241-242.

Schönemann, P. H. (1987). Jensen's *g*: Outmoded theories and unconquered frontiers, In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 311-327). New York: Falmer Press.

Schönemann, P. H. (1990). Facts, fictions, and common sense about factors and components. *Multivariate Behavioral Research, 25* (1), 47-51.

Schönemann, P. H. (1992). Extension of Guttman's result from *g* to PC1. *Multivariate Behavioral Research, 27* (2), 219-224.

Schönemann, P. H. (1997). Famous artefacts: Spearman's hypothesis. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 16* (6), 665-694.

Segre, E. (1969). *Great men of physics: The humanistic element in scientific work.* Los Angeles, CA: Tinnon-Brown.