Personality and Individual Differences 50 (2011) 196–200

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Exploratory factor analysis revealing complex structure ☆

Suitbert Ertel *

Georg-Elias-Müller Institut für Psychologie der Universität Göttingen, Gossler Strasse 14, 37073 Göttingen, Germany

ARTICLE INFO

Article history:
Received 5 July 2010
Received in revised form 11 September 2010
Accepted 22 September 2010
Available online 25 October 2010

Keywords:
Exploratory factor analysis
Simple structure
Complex structure
Factorial stability
Interpretability
Decathlon
Athletics

ABSTRACT

The study introduces varimin, a novel factorial rotation which, unlike Thurstone's principle of simple structure, attempts to model complexity. Varimin-rotated factors are conceived as components of functional structure. Simple structure- (e.g., varimax-) rotated factors are conceived as representing indeterminate clusters of those components. An exploratory factor analysis was performed on decathlon scores from Olympic Games 1948–1988 of 233 decathletes. I expected that an interpretation of factors of transparent physical variables, modeled by complex structure, should outdo an interpretation of factors modeled by simple structure. Results of factor transformations by varimin and varimax were compared. Varimin factors of the 10 decathlon events pointed to components contributing jointly, with varying degrees, to the decathletes' performances revealing the following components, F_1: general athletic energy, F_2: pacing of energy expenditure (speed vs. endurance), and F_3: location of prime energy expenditure (upper vs. lower body parts). Varimax factors clustered the sports events without consistency, functional features of physical activities were not revealed. An analysis of complex structure is deemed appropriate to revive, on a broader scale, exploratory factorial research which, due to questionable output in the past, has long since lost its earlier challenge.

1. Introduction

Discontent with exploratory factor analysis prevails as a common thread in its long history. Harsh judgments were made by Revelle (1983): "*Factors are fictions*"; by Eysenck (1992): "*Morass of factor analysis*"; Schönemann (1996): "*Psychopathology of factor indeterminacy*." Disappointment has become evident, more noticeably than elsewhere, by factorial investigations into athletic performance (Büsch, Hagemann, & Thielke, 2001; Teipel, 1988, p. 341 ff.). The poor factorial validity of sports motor tests has been summarized by Bös (1987, p. 141). "*Doubts regarding the acceptability of factorial results exist since long.*" (p. 461). Unfortunately, researchers neglected to ask why conventional procedures did not meet initial expectations despite decade-long application.

I did make an attempt at solving the problem (Ertel, 2009a, in press). It seemed to me that the principle of simple structure, introduced by Thurstone (1935, 1947) as a guideline for factor rotation, was the main cause of flawed factorial results. The aim of transforming factors to simple structure has almost never been questioned. Rather it has been regarded, ever since Thurstone's introduction, as self-evident and thus comparable to Lakatos' hard core of suppositions.

At this point, Thurstone's parsimony needs reconsideration. His principle ignores the fact that manifest variables, multifaceted as they are, generally arise by joint contributions of co-variance sources. Units of observation are generally engendered by functional interactions. Thurstone's mathematical principle lacks, what might be called, 'combinational prudence'. The mathematical simplicity of simple structure, destroying factorial combinations, is imposed, tacitly and blind, on seemingly solitary observational entities ("variables") while the underlying components of these entities are entirely ignored. Simple structure rotation forces variables into clusters while the sources of clustering remain obscure. Empirical research demands an unveiling of relations among functional components, but this demand is obstructed by Thurstone's doubtful methodical decision.

Consequently, I replaced the standard procedure of factor rotation, varimax, with varimin, a novel procedure for rotating factor coordinates with the aim of letting manifest variables (items, test scores, etc.) be loaded with as many factors as the co-variance data permit. Varimin searches and maximizes the initial complexity of extracted factors. Varimax rotation maximizes the variance of squared factor loadings across variables ("the varimax criterion"). Varimin does the opposite, the variance of squared factor loadings is minimized. Varimin thus aims at modeling complex structure, the antithesis to simple structure. (See Harman, 1968, p. 144, his formula 14.25 represents the Varimax criterion which is not

☆ A longer German version of this paper and with confirming results of a questionnaire presented to decathletes may be obtained on request.

* Address: Tobias-Mayer-Weg 3, 37077 Göttingen, Germany. Tel.: +49 551 2810251.

E-mail address: sertel@uni-goettingen.de

0191-8869/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2010.09.026

altered except that the goal of maximizing V is replaced with that goal of minimizing V.)

$$V = n \sum_{p=1}^{m} \sum_{j=1}^{n} (b_{jp}/h_j)^4 - \sum_{p=1}^{m} \left(\sum_{j=1}^{n} (b_{jp}^2/h_j^2)^2 \right)$$

At first sight, this approach might appear maverick (for an elaborate discussion see Ertel, 2009a, 2009b). Simple structure transformations, Kaiser's beloved "Little Jiffy", can hardly be abolished by theoretical reasoning alone. Appropriate empirical results are required in order to find out whether varimin transformations are not merely admissible, they might be better than varimax transformations or even be indispensable. Previous results of varimin applications on personality traits (Ertel, unpublished) and intelligence test performance (Ertel, in press) are encouraging.

An appropriate domain for testing the new methodological approach is athletics. Factors of motor behavior, performed under common rules, should be easier to comprehend than factors obtained from, say, intelligence test performance. In what follows, an account of factorial analyses of decathlon data is given which should merely exemplify the kind of results expectable for data subjected to the new procedure.

In a decathlon, each athlete competes in 10 events whose performance is distributed over 2 days: Various physical capacities and skills are required for winning scores on day 1 in 100 m race, long jump, shot put, high jump, 400 m race, and on day 2 in 110 m hurdles, discus, pole vault, javelin, and 1500 m race. Point scores are awarded on the basis of times (track events) and distance (field events). The athletes' scores correlate among each other (Table 1), factor analysis is expected to uncover considerably less than 10 latent determinants.

Previous factorial analyses of decathlon data.

Two earlier factor analyses of decathlon data were found, both aiming at simple structure. Karvonen and Niemi (1953) analyzed data of 62 decathlon competitors of the Olympic Games 1938, 1948, and 1952. A "multiple factor analysis" was applied "according to Thurstone". "*The rotations were performed graphically*" (p. 129). More methodical details are not conveyed. Poor results of earlier factorial studies on athletic performance might have discouraged researchers to continue factor analyses in this field. Despite published decathlon data from Olympic Games and world-wide interest in international sports competition I did not unearth, since 1977, from our literature any report on factor analyses of decathlon data.

Individual differences among decathletes should reveal one general physical component ('g') in the first place, just as one general intellectual component is almost always revealed from individual differences among intellectual performers. Overall genetic physical predispositions as well as training histories vary among athletes and should become manifest by one broad source of variance. A rotation to complex structure should reveal, in addition, 'g'-modifying factors which should specify particular physical functions.

Complex structure modeling hardly encounters the problems as they usually arise with simple structure. A varimax rotation – to take the most popular procedure for analyzing intellectual performance – removes 'g', although 'g' announces itself as an initial (unrotated) factor (Jensen, 1998). Varimax rotation dissolves the initial first factor, its loadings are redistributed among seemingly independent factorial "primaries". Highmore and Taylor (1954) lament factorial results of sports data: "*...the basic factor, representing general athletic ability (in which we are primarily interested), necessarily disappears, and the group factors [of simple structure rotation] show little relation to the classification indicated by the [initial] bipolar matrix*" (p. 4).

Now, since 'g' cannot be dismissed, theoretically, it must eventually be recovered from those "primaries". A "reunification" of variance is performed on a so-called second order level – a last-minute repair, as it were, with the help of artful mathematical operations (Schmid-Leiman transformation). A simple question arises which should have bothered statisticians since long: Why should 'g', present with an initial factorial solution, be removed at all? Varimin, by contrast, preserves 'g', improves its pattern together with patterns of additional factors while its contribution to the total variance is slightly diminished (to be demonstrated below). Additional factorial components will be aligned by varimin on the same first order level of factorial representation. They should emerge by itself without arbitrarily and riskily modeling them in advance.

2. Method

2.1. Data

Four sources of decathlon data were found and used: Intercorrelation matrices in Linden (1977), accessible by Basilevsky (1994), as well as individual scores for the 10 events from an internet source (2004), data from Kunz (1980) and from Zarnowski (1989). The largest number of athletes is provided by Zarnowski ($N = 233$), his athletes participated at 11 Olympic Games (1948–1988). The diversity within the sample is also the largest (49 participating nationalities). Zarnowski's data were therefore used for our factorial analysis.

2.2. Procedure

The 10 Zarnowski decathlon variables were intercorrelated. The intercorrelation matrix was subjected to Principal Component

Table 1
Intercorrelations of decathlon scores, taken from Zarnowski.

	1 100 m-run	2 Long jump	3 Shot put	4 High jump	5 400 m-run	6 110 m hurdles	7 Discus throw	8 Javelin throw	9 Pole vault	10 1500 m-run
1	1.00									
2	.66	1.00								
3	.51	.56	1.00							
4	.45	.64	.50	1.00						
5	.66	.60	.39	.54	1.00					
6	.62	.67	.55	.63	.58	1.00				
7	.43	.48	.80	.44	.37	.49	1.00			
8	.48	.61	.56	.70	.59	.63	.52	1.00		
9	.34	.46	.59	.40	.44	.41	.51	.52	1.00	
10	.09	.21	.09	.35	.53	.21	.12	.37	.24	1.00

Note: Source Zarnowski (1989): performances at the Olympic Games 1948–1988.
Negative time measures were used to obtain positive achievement correlations.