

Viktorija Trninić¹

Igor Jelaska¹

Janez Štalec²

796.01:159.923

Review paper

¹Faculty of Kinesiology, University of Split, 21000 Split, Croatia

²Faculty of natural sciences mathematics, University of Zagreb, 10020 Zagreb, Croatia

APPROPRIATENESS AND LIMITATIONS OF FACTOR ANALYSIS METHODS UTILISED IN PSYCHOLOGY AND KINESIOLOGY- PART 2

Abstract

Structural modelling techniques and application of models that extract latent variables are recent predominant techniques in the applied multivariate statistical procedures in social sciences. We believe that correlation studies can provide adequate findings if they are supported by logical analysis or causal modelling procedures. It is important to emphasize that factor analysis methods alone do not reveal the cause of covariability and that the final result of factor analytical investigation depends, in part, on the decisions and interpretations of the researcher. The question of the minimum sample size in factor analysis, ambiguousness of results obtained by FA and mathematical problems in the use of FA is particularly scrupulously discussed in this paper. Furthermore, the effect of the factor analysis of data obtained from experiments on the scientific paradigm was analyzed, with emphasis on the current problems with its application in social sciences research. Neither method, including factor analysis, is sufficient to answer all problem issues in the field of psychology and kinesiology. Therefore, it is necessary to combine complementary methods within the research which will allow a more comprehensive analysis of the researched phenomena and a greater validity of empirical results. Finally, reducing theories in psychology to a psychometric method and theories in kinesiology to a kinesiometric method is an anomaly of numerous quantitative studies within these scientific disciplines, making identification, instead of explanation of multi-causal nature of psychological and kinesiological phenomena, a primary focus of the research.

Key words: FACTOR ANALYSIS / PARADIGM / SAMPLE / AMBIGUOUSNESS OF RESULTS / MATHEMATICAL PROBLEMS

INTRODUCTION

Establishing relations between the phenomena and the cause of those phenomena, as well as defining the multicausal model of personality are the fundamental goals of personality psychology, while determining optimum anthropological patterns in a sport is one of the basic goals of sports kinesiology (Trninić, 1995).

Factor analysis has been used in numerous areas. For example, it is used in research studies deal-

ing with the structure of personality, intelligence, motor abilities, morphological characteristics and other areas (Momirović, 1966a, 1969, 1971, 1972; Momirović, et al., 1984), that is, wherever numerous correlated variables occur simultaneously and where the aim is to determine fundamental sources of covariates among the researched data.

Metikoš, Gredelj and Momirović (1979) state that the classical theory of motor abilities as latent dimensions can be abandoned by substituting factor

models with taxonomic models. These authors also state that taxonomic reasoning is not only empirically acceptable; it provides adequate results in the areas of its use.

In this context, Malacko and Popović (2001) emphasize that none of the theoretical structural models tested, based on the intercorrelation of some manifest variables cannot be accepted unless confirmed by confirmative factor analysis techniques. Suhr (2003) states that FA helps an investigator to determine the number of latent constructs underlying a set of items (variables), provides a means of explaining variation among variables (items) using a few newly created variables (factors), thus condensing information to define the content or meaning of factors, that is, latent constructs. The author listed the following assumptions underlying scientific exactness of EFA: interval or ratio level of measurement; random sampling; relationship between observed variables is linear; a normal distribution (of each observed variable); a bivariate normal distribution (each pair of the observed variables), and multivariate normality.

Furthermore, Costello and Osborne (2005) see FA as a complex and multi-step process. Suhr (2003) indicates that CFA allows the researcher to test the hypothesis that a relationship between the observed variables and their underlying latent construct(s) exists. In doing so, the researcher uses knowledge of the theory, empirical research, or both, postulates the relationship pattern a priori and then tests the hypothesis statistically.

Within contemporary psychology and kinesiology FA methods have proved useful in all structures where numerous correlated variables occur simultaneously in a research and where the aim is to determine fundamental sources of covariance among the data. Observation of these interrelated phenomena is particularly important in various fields of psychology, sociology, pedagogy, political science, economics, anthropology and medicine but also in chemistry, pharmacology and other sciences.

In kinesiology of sport, for example, the scientific observation of motor behaviour of athletes may reveal that many of them achieve similar results in numerous different motor activities. Therefore, in motor activities such as jumps, throws, sprints and direction changes, it would be feasible to presume the

existence of certain common factors which determine the intensity and successfulness of athlete's performance. These common factors which determine the successfulness of motor behaviour represent motor abilities. Motor abilities are hypothetical constructs, meaning they cannot be directly observed, and therefore, cannot be measured. Therefore, they can only be assessed indirectly, that is why they are called latent dimensions (Dizdar, 2006). According to Marsh (2007), all constructs in sport and exercise psychology are hypothetical constructs, and so must be validated using a construct validity approach. Construct validation is relevant to experimental as well as non-experimental research and is useful in evaluating the researcher's interpretation of the experimental manipulation.

Pervin, Cervone and John (2005) consider factor analysis not quite unbiased. Furthermore, Cervone and Pervin (2008) state that this statistical procedure identifies patterns of covariation in test responses but it does not answer the question of why the responses covary. It is the researcher, using his or her knowledge of psychology and kinesiology and relying on his or her theoretical beliefs, who infers the existence of some common entity (the factor) and interprets it. Different psychologists may make different interpretations.

Accordingly, Pervin, Cervone and John (2005) say the ultimate result of factor analytical research partially depends upon the decisions and interpretations of the researcher. Ergo, if the main goal of science is to determine laws of interrelations among the phenomena in nature and society as well as to investigate and conclude about their causes (Mejovšek, 2008) then FA, as a multivariate method, has a certain function in determining basic latent dimensions of a particular space. Reise, Weller and Comrey (2000) also state that the goal of factor extraction is to identify the number of latent dimensions (factors) needed to accurately account for the common variance among the items.

Fulgosi (1984) argues that by FA many useful constructs have been identified and many theories and models have been tested. Also, many wrong interpretations and convictions have been rejected, which, if it wasn't for FA, might have been trusted due to their new name, logical consistency, or similar

reasons. He also states that FA has been successfully used primarily in psychology: psychology of intelligence, personality, psychomotorics, physiological psychology, psychology of memory, social psychology, industrial psychology, educational psychology, and clinical psychology, as well as in other social and humanistic scientific fields.

Quite opposite is the opinion of MacCallum and Tait (1986). They analysed 152 scientific articles, published in the *Journal of Applied Psychology, Organizational Behavior and Human Performance* and *Personnel Psychology* in the period 1975-1984. They concluded that the selection of FA methods was frequently inappropriate, by which basic concepts of research methods and scientific work had been jeopardized. Furthermore, comparison of results by intervals (1975-1979, 1980-1984) made a minimal difference in the selections incurred.

We believe that, in the last decade or so, we can find more recent methodological approaches as, for example, Structural Equation Modeling (SEM) techniques/methods, tending to be prevailing methodological tools used for analysis and extraction of relations and latent relations between variables in psychology and anthropology (Loehlin, 2003; Kline, 2005; Hox, & Roberts, 2010).

It is important to point out that SEM is one of the appropriate multivariate methods which enable testing the theory and determining causal relations between phenomena.

Furthermore, SEM can be defined as a set of methods used to hypothesize about arithmetic means, variances and covariances of empirically obtained data in terms of smaller number of "structural" parameters defined by a hypothetical model (Kaplan, 2000). SEM is based on CFA, regression analysis and other multivariate methods (Marsch, 2007; Mejovšek, 2008; Milas, 2009). Milas (2009) states that SEM is a broad and general statistical approach, used to verify assumptions about relations between observed and latent variables. It can be used to test quantitatively defined theoretical models and to establish their empirical sustainability (Milas, 2009). In addition, SEM can be used to determine the structure of variables which describe a certain construct, phenomena or process and define their hypothetical causal relation. Steps in the implementation of SEM

include: determination (specification of the model), assessment and establishing suitability of the model, and modification and interpretation of the model (Milas, 2009).

In accordance with the above mentioned, Schutz and Gessaroli (1993) argued that CFA and SEM were the most important statistical tools for the 1990s, but that those techniques had seen almost no application in sport and exercise psychology.

Marsh (2007) claimed that CFA/SEM should be the methodology of choice and recommended that sport and exercise psychologists make greater use of these techniques.

Furthermore, Marsh (2007) points out that despite this growing popularity, there appears to be an ever-widening gap between state-of-the-art methodological and statistical techniques that should be part of the repertoire of serious empirical researchers and the actual skill levels of many applied researchers.

The same author states that, because CFA/SEM has considerable flexibility for addressing complex substantive issues in sport and exercise psychology, there is an increasing need for heuristic, nontechnical overviews of these techniques to assist potential users.

Also, Marsh (2007) suggests that there is a wide variety of multivariate quantitative analyses (CFA, SEM, causal modelling, multilevel analysis, longitudinal data analysis and multitrait-multimethod designs) to address substantive and theoretical issues from a construct validity perspective.

Finally, Marsh (2007) claims that CFA/SEM as statistical techniques provide new and interesting ways of resolving existing problems in sports psychology. The author also points out that SEM technology is more accessible than ever to applied researchers via modern software programs.

FACTOR ANALYSIS AND SCIENTIFIC PARADIGMS

Kuhn (1970, 1974) noticed, especially in the area of social sciences, disagreement in the nature of legitimate scientific problems and methods, so he even used the term: *normal science* (1970, 1999), inherently meaning science performing under the influence of the current, prevalent paradigm. When science can no longer explain aberrations, which are be-

coming ever deeper through time, a crisis in science occurs followed by a resolving scientific revolution and by the establishment of a new paradigm, a new foundation of normal science.

Gorsuch (1990, 1997) emphasized a reciprocal influence between research methods of standard use and currently accepted paradigm. Furthermore, the author noted that, under the influence of different scientific paradigms, the researcher will often choose a different method of factor analysis. Paradigms often influence what techniques are used, in addition to the logic of those techniques, and history affects paradigms. The same author also assumes that different paradigms and purposes will indeed lead to different factor analysis (Gorsuch, 1990, 1997).

In accordance with the aforementioned, Mejovšek (2008) declared that in natural sciences one of the crucial arguments for the new paradigms acceptance and the already existent paradigms rejections was a greater quantitative exactness of the new paradigms and their greater predictive power or precision of future events. A researcher's approach to the investigation of certain aspects of particular scientific issues will depend on a paradigm he/she prefers. The author noticed that the greater exactness has been verified and validated by means of FA methods, meaning FA is a tool for the validation of the existent paradigm. Also, with the development of science and scientific research intensively conducted in a particular area, the existing paradigms are evaluated and those paradigms are retained that can best meet the requirements of the scientific demands in revealing the laws by which phenomena in a particular area occur.

Therefore, Mejovšek (2003) states that the more complex a certain area is, the more probable is the existence of a larger number of paradigms. This is the evident situation in psychology and kinesiology, in the opinion of the authors of this paper.

Methodology used to identify the structure of personality traits, factor analysis, is often challenged for not having a universally-recognised basis for choosing among solutions with different numbers of factors. Thus for example, a five factor solution depends to some degree on the interpretation of the analyst. This has led to disputes about the „true“ number of factors (Pervin, Cevone & John, 2005).

In psychology, for example, no unique theory of personality exists. There is no unique theory of motor abilities in kinesiology either. It can be concluded that FA generates the whole semantic continuum and a researcher interprets and infers. Fornell (1987, in Bucik, 1990) states that multivariate methods “of the second generation” (redundancy analysis, covariate structures analysis, confirmatory multidimensional scaling) are more rigorous than those “of first generation” (MANOVA, ANCOVA, multiple regression,...) regarding establishing theoretical starting points, so they are verified by confirmatory techniques (Vodopivec, 1988).

For Pervin, Cervone, and John (2005) personality traits are hierarchically organized. Momirović and associates (1987) explain that the hypothetic model of motor abilities is also of a hierarchical type, that is, it consists of one general factor and lower structures of lower-order factors. Such a concept of human characteristics and abilities can be a foundation for a scientific theory. In psychology, for example, the hierarchical standpoint, established by Eysenck (1980a, 1980b), suggests that at the simplest level behaviour may be observed through specific responses (the so called specific response level). Some of these responses are becoming associated and form more general habits (the so called habitual response level). At a higher level of hierarchical organization different personality traits are associated and form higher-order factors or super-factors (Eysenck, 1970, 1990).

The second most influential psychologist of the 20th century, Cattell (1965) preferred to work with a large number of factors in research of personality traits. As opposed to him, Eysenck and Eysenck (2005) used FA to combine personality traits into a smaller number of uncorrelated superfactors, which cover a broader range of behaviour.

The major difference between Eysenck and Cattell is that Cattell prefers to work with larger number of factors at the trait level, which have a more narrow definition but tend to correlate with each other. In contrast, Eysenck uses secondary FA to combine traits into a smaller number of superfactors, which cover a broader range of behaviour and tend to be uncorrelated.

SAMPLE SIZE IN FA

In studies treating FA, considerable attention has been paid to the problem of sample size. Typical suggestions embrace information on the minimum sample sizes and the minimum ratio between sample size (N) and the number of manifest variables (p), that is, the minimum ratio $N : p$.

Analysis of problems of sample size was interpreted by Kline (1979) who recommended that N should be at least 100, while different authors argued that N should be at least 200, and Cattell (1978) claimed the minimum desirable N to be 250. Comrey and Lee (1992) urged researchers to obtain samples of 500 or more observations whenever possible in factor analytic studies.

Tinsley and Tinsley (1987) also reported a problem of sample size in FA: the sample of up to 100 entities is poor, up to 200 entities fair, up to 300 entities good, up to 500 entities very good, and up to 1000 entities excellent.

Considering recommendations for the $N : p$ ratio, Cattell (1978) believed it should be in the range of 3 to 6. Gorsuch (1983) argued for a minimum ratio of 5. Everitt (1975) recommended that the $N : p$ ratio should be at least 10. MacCallum, et al. (1999) commented the wide range in these recommendations. Furthermore, the same authors have noticed that the inconsistency in the recommendations can probably be attributed partly to a relatively small amount of explicit evidence or support provided for any of them. Most of them seem to be general guidelines developed from substantial experience on the part of their supporters. Some authors (e.g., Comrey, & Lee, 1992) placed the sample size question into the context of the need to make standard errors of correlation coefficients adequately small so that following FA of those correlations would yield stable solutions.

There are some research findings that are relevant to the sample size question. There is considerable literature, for instance, on the topic of standard errors in factor loadings. As sample size increases, the variability in factor loadings across repeated samples will decrease. Formulas for estimating standard errors of factor loadings have been developed for various types of unrelated loadings (Girshick, 1939; Jennrich, 1974; Lawley, & Maxwell, 1971) and rotated loadings (Archer, & Jennrich, 1973; Jennrich, 1973, 1974).

Cudeck and O'Dell (1994) reviewed the literature on this subject and provided further developments for obtaining standard errors of rotated factor loadings. Application of these methods would demonstrate that standard errors decrease as N increases. In addition, Archer and Jennrich (1976) showed this effect in the Monte Carlo study. Although this effect is well defined theoretically and has been demonstrated with simulations, there is no guidance available to indicate how large N must be to obtain adequately small standard errors of loadings. A detailed study of this question would be difficult because, as emphasized by Cudeck and O'Dell (1994), these standard errors depend in a complex way on many things other than sample size, including method of rotation, number of factors, and the degree of correlation among the factors. Furthermore, a general method for obtaining standard errors for rotated loadings has been developed (Browne, & Cudeck, 1997).

MacCallum, et al. (1999) state that the factor analysis literature includes a range of recommendations regarding the minimum sample size necessary to obtain factor solutions that are adequately stable and that correspond closely to population factors. Also, a fundamental misconception about this issue is that the minimum sample size, or the minimum ratio of sample size to the number of variables, is invariant across studies. Furthermore, the same authors claim that, in fact, necessary sample size is dependent on several aspects of any given study, including the level of communality of the variables and the level of overdetermination of the factors. In conclusion, the authors present a theoretical and mathematical framework that provides a basis for understanding and predicting these effects. The hypothesized effects are verified by a sampling study using artificial data. Results demonstrate the lack of validity of common rules of thumb and provide a basis for establishing guidelines for sample size in FA.

Accordingly, Marsh, et al. (1998), for the number of indicators (p) per factor (p/f) in CFA, by varying sample size ($N=50-1000$) and p/f (2-12 item per factor) in 35,000 Monte Carlo simulations, conclude that researchers should consider more indicators per factor than is evident in current practice.

MacCallum, et al. (1999), however, also claim, regarding sample size, that N is in fact highly depen-

dent on several specific aspects of a given study. Under certain conditions, a relatively small sample may be entirely adequate, whereas under other conditions, a very large sample may be inadequate. So, when discussing sample size, one must identify aspects of a study that influence the necessary sample size. This is first done in a theoretical framework and is then verified and investigated further with simulated data. On the basis of these findings MacCallum, et al. (1999) provided guidelines to estimate the necessary sample size in an empirical study. The theoretical framework they have presented provides a foundation for the following hypotheses about effects of sample size in FA.

We will cite their three crucial points:

“1. As N increases, sampling error will be reduced, and sample FA solutions will be more stable and will more accurately recover the true population structure.

2. Quality of FA solutions will improve as communalities increase. In addition, as communalities increase, the influence of sample size on quality of solutions will decline. When communalities are all high, sample size will have a relatively little impact on quality of solutions, meaning that accurate recovery of population solutions may be obtained using a fairly small sample. However, when communalities are low, the role of sample size becomes much more important and will have a greater impact on quality of solutions.

3. Quality of FA solutions will improve as overdetermination of factors improves. This effect will be reduced as communalities increase and may also interact with sample size.” (MacCallum, et al., 1999)

In their Monte Carlo study (MacCallum, et al., 1999), they retained the known correct number of factors in the analysis of each sample. Thus it was found that excellent recovery of population factors could be achieved with small samples under conditions of high communalities and optimal overdetermination of factors. However, an open question remains whether analysis of sample data under such conditions would consistently yield a correct decision about the number of factors. “We expect that this would, in fact, be the case simply because it would seem contradictory to find the number of factors to be

highly ambiguous but recovery of population factors to be very good if the correct number were retained. Nevertheless, if this were not the case, our results regarding recovery of population factors under such conditions might be overly optimistic because of potential difficulty in identifying the proper number of factors to retain. One other limitation of our Monte Carlo design involves the limited range of levels of N , p and r that were investigated. Strictly speaking, one must be cautious in extrapolating our results beyond the ranges studied. However, it must be recognized that our findings were supported by a formal theoretical framework that was not subject to a restricted range of these aspects of design, thus lending credence to the notion that the observed trends are probably valid beyond the conditions investigated in the Monte Carlo study. Our approach to studying the sample size question in factor analysis has focused on the particular objective of obtaining solutions that are adequately stable and congruent with population factors. A sample size that is sufficient to achieve that objective might be somewhat different from one that is sufficient to achieve some other objective, such as a specified level of power for a test of model fit (MacCallum, Browne, & Sugawara, 1996) or standard errors of factor loading - that are adequately small by some definition. An interesting topic for future study would be the consistency of sample sizes required to meet these various objectives.”

It is impossible to derive a minimum sample size that is appropriate in all situations (MacCallum, & Tucker, 1991). By using theoretical arguments and empirical evidence, these authors demonstrated that the minimum sample size needed to accurately recover a population factor pattern is a function of several variables including the variables-to-factor-ratio, the average communality of the variables, and the degree to which the factors are overdetermined (defined, in part, by the number of variables that load on each factor). Reise, Waller, and Comrey (2000) assert when communalities are high ($>.6$) and the factors are well defined (have many large loadings), sample sizes of 100 are often adequate. However, when communalities are low (e.g., when analyzing items), the number of factors is large and the number of indicators per factor is small, even a sample size of 500 subjects may not be adequate.

In addition to sample size, the second issue that warrants consideration is sample heterogeneity. In terms of identifying replicable factors, researchers should assemble samples with sufficient examinee representation at all levels of the trait dimensions (in order to accurately estimate the population item intercorrelations). One consequence of this rule is that using the standard pool of undergraduates may be suitable when undergraduates manifest sufficient heterogeneity with respect to trait standing. On some constructs, such as extraversion or agreeableness, this seems reasonable. For other constructs, however, such as uni-polar depression or psychotic ideation, undergraduates may not be an appropriate respondent pool to accurately map the factor space of clinical assessment scales.

Costello and Osborne (2005) state that the researchers using large samples and making informed choices from the options available for data analysis are the ones most likely to accomplish their goal: to come to conclusions that will generalize beyond a particular sample to either another sample or to the population (or *a* population) of interest. To do less is to arrive at conclusions that are unlikely to be of any use or interest beyond that sample and that analysis.

Comrey (1978) and Reise, Weller and Comrey (2000) indicate a danger if too few factors are extracted - a researcher may miss important distinctions among the items, and the subsequently rotated solution may be distorted in non-systematic ways. However, if too many dimensions are retained, some rotated factors may be ill defined with only one or two salient loadings. Comrey (1967) suggests an analytic rotation strategy specifically designed to address these issues. Yet, although there are many rules of thumb and statistical indices for addressing the dimensionality issue in EFA, no procedure seems entirely satisfactory. Fava and Velicer (1992a) as well as Wood, Tataryn and Gorsuch (1996) empirically verify the effects of under - and overextraction on the factor recovery of known population structures. They generally agree that it is preferable to extract too many factors rather than too few. For instance, on the basis of the highly ambitious Monte Carlo study, Tataryn and Gorsuch (1996) (p. 354) recently concluded that "(a) when underextraction occurs, the estimated factors are likely to contain considerable

error; [and] (b) when overextraction occurs, the estimated loadings for the true factors usually contain substantially less error than in the case of underextraction."

AMBIGUOUSNESS OF RESULTS OBTAINED BY FA

Cattell (Pervin, Cervone, & John, 2008) claims that if multivariate, factor analytic research is at all apt to determine the basic personality structure, then the same factors or personality traits should be obtained from all the three types of data: from life information, from the questionnaire data and from the objective test data. Eysenck (1991) also defines a criterion as a possibility to replicate results. That external measure is essential and condition *sine qua non* for taxonomic paradigm of a personality description system. Any problem connected with replicability automatically excludes that system from any further consideration.

The fact that different criteria produce different solutions, that is, proclaims different number of principal components to be significant is probably the greatest issue of FA. For example, the Guttman-Kaiser's criterion shows tendency to hyperfactorization - too many principal components become significant; whereas the PB criterion tends to hypofactorization - too few principal components become significant (Viskić-Štalec, 1991). If FA is so powerful mathematical-statistical procedure, as suggested by its advocates, then the same factors should be found in equivalent research (Pervin, Cervone, & John, 2005). Personality psychology researchers claim the five-factor model to be the fundamental finding and those five factors is "just the right number of factors" (McCrae, & John, 1992). Other researchers insist that less than five factors are enough (Eysenck, & Eysenck, 1985; Eysenck, 1983, 1991, 1993; Zuckerman, 1990). Quite opposite, Benet and Waller (1995), Buss (1988), Cattell (1990), Tellegen (1993) suggest five factors that generate manifest personality structure are not sufficient to describe personality.

Despite the consensus in relation to the Big-five model, there is a debate considering its status (Goldberg, & Saulcier, 1995; Pervin, 1994; Eysenck, 1991, 1993). Pervin and John (1997) have noticed that many critics suggest the level of congruence

among the conclusions of different studies has been less than ideal.

A number of critics suggest that the degree of correspondence among the findings from different studies has been less than ideal. In the words of one supporter of trait theory, „the resemblance is more fraternal than identical“ (Briggs, 1989, p. 248). In sum, we must question whether FA will provide the basic units of personality). Opposite to the prior attitudes, for Eysenck (1985), FA is the only and the best procedure for the detection of either basic or latent personality traits. As Bucik, Boben and Kranjc (1997) have asserted FA was in the past a powerful, effective mathematical method tool for the determination of contents of the personality constructs structure. However, even among the researchers who agree mostly about the methodological approach, there are still certain disagreements regarding the number of factors describing the personality structure.

Furthermore, Mlačić and Knezović (1997) report about two directions or currents of researchers in the interpretation of factors obtained by FA. Their opinion is that Cattell's position is most realistic: he explicitly equates the obtained factors with the neuro-psychological structures generating behavioural patterns. On the other end of the FA interpreters are researchers who utilize FA as a substitution for cluster analysis and do not make any presumption that the obtained factors may have any significance outside the investigated group of data (Goldberg, & Digman, 1994).

Mathematical problems with FA methods application

It must be noted that essential, mathematical problems occurring in application of FA methods are seldom discussed among kinesiologists. The authors of the present article reason that arithmetic restrictions of computer, which generates numerical problems, and limitations of mathematical statistic apparatus in general are restrictions that should be studied more thoroughly even in kinesiology and psychology. Also, it should be emphasized here that the generation of exact, differentially weighted (pondered) linear combination in practical research is an exception and that in the final computer arithmetic most data are approximations of a small relative error.

Furthermore, standard criteria aiming at assessing quality of certain methods of FA are based on the amount of the explained variance of particular factors. Our reasoning is that the further evaluation of FA methods should be based upon the detection of the distribution of test-statistics behaviour of certain criteria. A proposed approach might incite the usage of FA methods based upon a certain level of significance, that is, as statistical methods.

Bartlett (1950) has developed his famous test for the statistical significance of a correlation matrix, that is, the test of equivalency of all population's roots. Furthermore, Lužar (1983, 1984) claims that significance tests generate too large number of components, therefore, a large number of criteria is developed, based upon non-statistical foundations and motivated by various logic reasoning. Based on the simulated experiment, the authoress suggests it is sensible to use only three criteria: GK (Guttman, & Kaiser, 1956), PB (Štalec, & Momirović, 1971) and CH (Momirović, & Zakrajšek, 1972).

CONCLUSION

The next issue in FA application is determining the level of correlation among the factors. Furthermore, the orthogonal solutions only define the dimensionality of latent space, and oblique solutions, under conditions of simple structure, can mean more variance interpreted. It is clear that uniform coverage of manifest space enables more appropriate, yet still subjective explanations.

Despite the fact that FA is one of the main multivariate analyses, used both in theoretical and applied psychological and kinesiological research, there are serious disputes regarding its appropriateness and limitations. The arguments are listed as follows:

- The reality of existence of latent, initially hypothetic factors being responsible for a number of correlations among the observed variables. The factors are a probable mathematical-statistical referent framework, but it does not mean that the determined latent dimensions actually exist in the reality.

- Incorrect application of FA methods may lead to wrong conclusions.
- Known structures of factors are frequently not corroborated by empirical research, that is, the same factors have not been found in different empirical factorial findings.
- Insufficient definition of factors – expert and research teams in a specific field should be careful to meaningfully and coherently denote (name) and explain the obtained factors; a conflict between exactness and practicality may occur.
- The existence of many various rotations (because there are no quantitative indices about quality of an individual rotations), which generates semantic continuum of possibilities. Certain rotations are weighted based on the impression that they are more important than the others. A researcher chooses appropriate structure based on her/his subjective assessment. In doing so, her/his goal is to achieve as sensible interpretation of the extracted factors as possible.
- Sample size in FA – it depends primarily on the specific issues treated in a particular research study; therefore, it can hardly be generalized. Recommendations of most authors that sample should be twice or three times the number of variables is not theoretically justified, and practically they have been proven inadequate.

Researchers on different populations of athletes using different measuring instruments obtain very different findings. However, it is crucial whether particular extracted factors get high level of empirical support or not. After all the obtained factors have been interpreted, the final verification is performed based on the checking factors' intercorrelations. If correlations among the factors deviate from the expected, then we can doubt about the correctness of factors' interpretation and the interpretation procedure should be repeated.

Sample size should be emphasized again, especially in its relation to population size for which we are detecting a latent structure. In FA researchers should meet both the conditions for the significance of multiple correlations and be very prudent when selecting sample size. Many authors recommend a sample to be larger than the twice or triple the number of variables. This recommendation is not theoretically justified and it has shown to be unsuitable in practice.

From the application point of view, FA can be statistical technique or method only when hypotheses are defined very precisely covering homogeneously the whole investigated space. On the other hand, explorative FA techniques are mathematical methods used to reduce space to a fewer dimensions which should allow researchers optimal utilization of other mathematical and statistical methods. It can be concluded that in current kinesiology too much significance is given to EFA in comparison with CFA.

Furthermore, theory, assessment and quantitative research of personality in sports are insufficient for answering numerous issues and problems in sport science, relating the explanation of internal and external determinants of athlete's performance and sports achievement. This is due to reducing theories in psychology to a psychometric method and theories in kinesiology to a kinesiometric method which is an anomaly of numerous quantitative studies, focusing on identification of the obtained results, instead of explanation of mechanisms underlying psychological and kinesiological phenomena. We believe that integrating different complementary approaches to research (e.g. nomothetic and idiographic, quantitative and qualitative) enhances the ability of comprehension, enables greater validity of research findings as well as their practical applicability. Finally, psychologists and kinesiologists are in need of science of discourse and science of intervention which is oriented towards the processes of change.

REFERENCES

1. Archer, J.O., & Jennrich, R.I. (1973). Standard errors for orthogonally rotated factor loadings. *Psychometrika*, 38, 581–592.
2. Archer, J.O., & Jennrich, R.I. (1976). A look, by simulation, at the validity of some asymptotic distribution results for rotated loadings. *Psychometrika*, 41, 537–541.
3. Bartlett, F.C. (1950). Subjective Judgments. *Nature*, 166(4232), 984–985.
4. Benet, V., & Waller, N.G. (1995). The Big Seven factor model of personality description: Evidence for its cross-cultural generality in Spanish sample. *Journal of Personality and Social Psychology*, 69, 701–718.
5. Briggs, S.R. (1989). The optimal level of measurement for personality constructs. In D.M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (246–260). New York: Springer Verlag.
6. Browne, M.W., & Cudeck, R. (1997). *A general approach to standard errors for rotated factor loadings*. Paper presented at the European Meeting of the Psychometric Society, Santiago de Compostelo, Spain.
7. Bucik, V. (1990). Faktorska ali clusterska analiza – katero tehniko uporabiti. [Factor or cluster analysis – which technique to use. In Slovenian.] *Anthropos*, 22(5/6), 253–259.
8. Bucik, V., Boben, D., & Krajnc, I. (1997). Vprašalnik BFQ in ocenjevalna lestvica BFO za merjenje «Velikih pet» faktorjev osebnosti: slovenska priredba. [Questionnaire BFQ and grading scale BFQ for Big Five trait factors measuring – Slovenian version. In Slovenian.] *Psihološka obzorja*, 6(4), 5–34.
9. Buss, A.H. (1989). Personality as traits. *American Psychologist*, 44, 1378–1388.
10. Cattell, R.B. (1965). *The scientific analysis of personality*. Baltimore: Penguin.
11. Cattell, R.B. (1978). *The scientific use of factor analysis*. New York: Plenum.
12. Cattell, R.B. (1990). Advances in Cattellian personality theory. In L.A. Pervin (Ed.), *Handbook of personality: Theory and research* (101–110). New York: Guilford Press.
13. Cervone, D., & Pervin, L.A. (2008). *Personality. Theory and research*. New York, John Wiley and Sons.
14. Comrey, A. L. (1967). Tandem criteria for analytic rotation in factor analysis. *Psychometrika*, 32(2), 143–154.
15. Comrey, A.L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46, 648–659.
16. Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.
17. Costello, A.B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7), 1–9.
18. Cudeck, R., & O'Dell, L.L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*, 115, 475–487.
19. Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet Sveučilišta u Zagrebu.
20. Everitt, S. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry*, 126, 237–240.
21. Eysenck, H.J. (1970). *The structure of personality*. (3rd ed.) London: Methuen.
22. Eysenck, H.J. (1980a). The bio-social nature of man. *Journal of Social and Biological Structures*, 1, 125–134.
23. Eysenck, H.J. (1980b). The bio-social model of man and unification of psychology. In A.J. Chapman & P.M. James (Eds.), *Models of man*, London: British Psychological Society.

24. Eysenck, H.J. (1983). Is there a paradigm in personality research? *Journal of Research in Personality*, 17, 369–397.
25. Eysenck, H.J., & Eysenck, M.W. (1985). *Personality and individual differences*. New York: Plenum.
26. Eysenck, H.J. (1990). Biological dimensions of personality. In A. Pervin (Ed.), *Handbook of personality: Theory and research* (244-276.). New York: Guilford Press.
27. Eysenck, H.J. (1991). Dimensions of personality. 16, 5 or 3? - Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12(8), 773–790.
28. Eysenck, H.J. (1993). Comment on Goldberg. *American Psychologist*, 48, 1299–1300.
29. Fava, J.L., & Velicer, W.F. (1992). The effects of overrelaxation on factor and component analysis. *Multivariate Behavioral Research*, 27, 387–415.
30. Ford, K., MacCallum, R., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291–314.
31. Fulgosi, A. (1984). *Faktorska analiza*. [Factor analysis. In Croatian-Serbian.] Zagreb: Školska knjiga.
32. Girshick, M.A. (1939). On the sampling theory or roots of determinantal equations. *Annals of Mathematical Statistics*, 10, 203–224.
33. Goldberg, L.W., & Digman, J.M. (1994). Revealing structure in the data: Principles of exploratory factor analysis. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (216–242). New York: Springer.
34. Goldberg, L.R., & Saulcier, G. (1995). So what do you propose we use instead? A reply to Block. *Psychological Bulletin*, 117, 221–225.
35. Gorsuch, R.L. (1983). *Factor analysis*. (2nd ed.) Hillsdale, NJ: Erlbaum.
36. Gorsuch, R.L. (1990.) Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, 25(1), 33–39.
37. Gorsuch, R.L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, 68(3), 532–560.
38. Hox, J., & Roberts, J.K. (2010). *Handbook of Advanced Multilevel Analysis*, Routledge Academic.
39. Jennrich, R.I. (1973). Standard errors for obliquely rotated factor loadings. *Psychometrika*, 38, 593–604.
40. Jennrich, R.I. (1974). Simplified formulae for standard errors in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology*, 27, 122–131.
41. Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks: Sage.
42. Kline, P. (1979). *Psychometrics and psychology*. London: Academic Press.
43. Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. The Guilford Press.
44. Kuhn, T.S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
45. Kuhn, T.S. (1974). Second thoughts paradigms. In F. Suppe (Ed.), *The structure of scientific theories* (459–482). London: University of Illinois Press.
46. Kuhn, T.S. (1999). *Struktura znanstvenih revolucija*. Zagreb: Naklada Jesenski i Turk, Hrvatsko sociološko društvo.
47. Lawley, D.N., & Maxwell, A.E. (1971). *Factor analysis as a statistical method*. New York: Macmillan.
48. Loehlin, J.C. (2004). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*, (4th ed). Mahwah, NJ: Lawrence Erlbaum Associates.
49. Lužar, V. (1983). Aproksimacije distribucija svojstvenih vrijednosti matrica u standardiziranom, Guttmanovom i Harrisovom prostoru. [Distributions' approximations of eigenvalues of matrices

- in standardized, Guttman and Harris space. In Croatian-Serbian.] (Doctoral dissertation, University of Zagreb) Zagreb: Elektrotehnički fakultet.
50. Lužar, V. (1984). Monte Carlo study of the distributions of the roots of a Harris matrix. In V. Lužar & M. Cvitaš (Eds.), Proceedings of 6th International Symposium "Computer at the University", Dubrovnik, (5111–5118). Zagreb: University Computing Centre.
51. MacCallum, R.C., Browne, M.W., Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149.
52. MacCallum, R.C., & Tucker, L.R (1991). Representing sources of error in the common factor model: Implications for theory and practice. *Psychological Bulletin*, 109, 502–511.
53. MacCallum, R.C., Widaman, K.F., Zhang, S.B., & Hong, S.H. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99.
54. Malacko, J., & Popović, D. (2001). *Metodologija kineziološko antropoloških istraživanja*. [Methodology of kinesiological anthropological research. In Serbian.] Leposavić: Fakultet za fizičku kulturu Univerziteta u Prištini.
55. Marsh, H.W., Hau, K., Balla, J.R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33(2), 181–220.
56. Marsh, H.W. (2007). Application of Confirmatory Factor Analysis and Structural Equation Modeling in Sport and Exercise Psychology. In G. Tenenbaum & R. C. Eklund (Ed.), Handbook of sport psychology (774–798). Hobelen, N.J.: Wiley.
57. McCrae, R.R., & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175–215.
58. Mejovšek, M. (2008). *Metode znanstvenog istraživanja u društvenim i humanističkim znanostima*. [Scientific research methods in social and humanistic sciences. In Croatian.] Jastrebar-sko: Naklada Slap.
59. Metikoš, D., Gredelj, M., & Momirović, K. (1979). Struktura motoričkih sposobnosti. [Structure of motor abilities. In In Croatian-Serbian.] *Kineziologija*, 9(1-2), 25–46.
60. Milas, G. (2009). Istraživačke metode u psihologiji i drugim društvenim znanostima. Jastrebarsko: Naklada Slap.
61. Mlačić, B., & Knezović, Z. (1997). Struktura i relacija Big Five markera i Eysenckova upitnika ličnosti: empirijska usporedba dvaju strukturalnih modela ličnosti. [The structure and relation of the Big Five markers and the Eysenck's personality questionnaire: An empirical comparison of two personality structural models. In Croatian.] *Društvena istraživanja*, 6(1), 1–27.
62. Momirović, K. (1966b). Valjanost psiholoških mjernih instrumenata. [Validity of psychological measuring instruments. In Croatian-Serbian.] In *Odabrana poglavlja iz psihometrije i neparametrijske statistike* (55–97). Zagreb: Društvo psihologa Hrvatske, Republički zavod za zapošljavanje.
63. Momirović, K. (1969). *Faktorska struktura antropometrijskih varijabli*. [Factor structure of anthropometric variables. In Croatian-Serbian.] Zagreb: Institut za kineziologiju, Visoka škola za fizičku kulturu.
64. Momirović, K. (1971). *Struktura i mjerenje patoloških konativnih faktora*. [The structure and measuring of pathological conative factors/personality traits. In Croatian-Serbian.] Zagreb: Republički zavod za zapošljavanje.
65. Momirović, K. (1972). *Metode za transformaciju i kondenzaciju kinezioloških informacija*. [Methods for the transformation and condensation of kinesiological information. In Croatian-Serbian.] Zagreb: Institut za kineziologiju, Fakultet za fizičku kulturu.
66. Momirović, K., Štalec J., Prot, F., Bosnar K., Pavičić, L., Viskić-Štalec, N., & Dobrić, V. (1984). *Kompjutorski programi za klasifikaciju, selekciju, programiranje i kontrolu treninga*. [Computer programmes for the classification, selection, and training programming and con-

- trol. In Croatian-Serbian.] Zagreb: Fakultet za fizičku kulturu.
67. Momirović, K., Prot, F., Dugić, D., Knezović, Z., Bosnar K., Erjavec, N., Gredelj, M., Kern, J., Dobrić, V., & Radaković, J. (1987). *Metode, algoritmi i programi za analizu kvantitativnih i kvalitativnih promjena*. [Methods and algorithms for the analysis of quantitative and qualitative transformations / changes. In Croatian-Serbian.] Zagreb: Institut za kineziologiju, Fakultet za fizičku kulturu.
68. Pervin, L.A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, 5, 103–113.
69. Pervin, L.A., & John, O.P. (1997). *Personality: Theory and research*. (7th ed.) New York: John Wiley & Sons.
70. Pervin, L.A., Cervone, D., & John, O.P. (2005). *Personality: Theory and research* (9th ed.) New York: John Wiley & Sons.
71. Pervin, L.A., Cervone, D., & John, O.P. (2008). *Psihologija ličnosti: Teorije i istraživanja*. Zagreb: Školska knjiga.
72. Reise, S.P., Waller, N.G., & Comrey, A.L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12(3), 287–297.
73. Schutz, R. W., & Gessaroli, M. E. (1993). *Use, misuse and disuse of psychometrics in sport psychology research*. In R. N. Singer, M. Murphy & L.K. Tennath (Eds.), *Handbook of research on sport psychology* (901–917). New York: Mac-Millan.
74. Suhr, D. (2003). Reliability, exploratory and confirmatory factor analysis for the scale of athletic priorities. In *Proceedings of the 28th Annual Meeting of SAS Users Group International*. Cary, NC: SAS Institute.
75. Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychology and Inquiry*, 4, 122–133.
76. Tinsley, H.E.A., & Tinsley, D.J. (1987). Use of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34, 414–424.
77. Trninić, S. (1995). *Strukturna analiza znanja u košarkaškoj igri*. [Structural analysis of knowledge in basketball game. In Croatian.] (Doctoral dissertation, University of Zagreb) Zagreb: Fakultet za fizičku kulturu Sveučilišta u Zagrebu.
78. Viskić-Štalec, N. (1991). *Elementi faktorske analize-skripta* [The elements of the factor analysis-script. In Croatian.] Zagreb: Fakultet za fizičku kulturu.
79. Vodopivec, B. (1988). Epistemološki vidiki analize kovariančnih struktur. [Epistemological perspectives of covariate structure analyses. In Croatian-Serbian]. *Anthropos*, 18, 46–57.
80. Wood, J.M., Tatryn, D.J., & Gorsuch, R.L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, 1, 354–365.
81. Zuckerman, M. (1990). The psychophysiology of sensation seeking. *Journal of Personality*, 58, 313–345.

The paper is a result of work within the framework of the research project number 315-0342607-1815 (Diagnostics of specific psychosocial features of athletes in team sports games) granted by the Ministry of Science, Education and Sport of the Republic of Croatia.

Received: 08.09.2012.
Accepted: 22.10.2012.