Review

On the factor extracted by factor analysis as abillity, such as physical fitness and motor ability

MATSUURA Yoshiyuki *

スポーツ科学において，測定される諸身体的属性には体力，運動能力，運動技能やその他の種々の能力及び身長，体重のように定義された属性がそのまま可視的である身体的属性がある。後者は多くの場合目に見える属性であり，その属性の定義に従って測定が行われ，測定値がそのままその属性の計量値である。しかし、前者の諸能力は目に見えず，直接測定のための尺度を当てがう事は出来ない。特に，能力の測定に運動成就テスト（performance test）を用いる場合は，その測定値は測定したい能力が発揮され成就された結果の計量値である。したがって，この成就結果の測定結果から能力を推定せざるを得ない。そのためには，能力の計量値と成就結果の計量値との関係が定立される必要がある。測定値が多くの場合数値で示される事から能力と成就結果の測定値との間に数学的関数関係が定立される事が推定には便利であり，比較的客観的推定が可能となる。成就結果と能力の関係を示す関数として項目特性関数が工夫されているが，因子分析の基本的仮定である測定値と因子に関する $X = \Lambda\Theta + aU$ は項目特性関数 $X = \phi(\theta) + e$ である。項目特性関数において，ϕ を一次関数とし，e を因子分析で云う特殊性（誤差を含む）とすれば，項目特性関数は因子分析における基本的仮定である。この事から，因子分析によって抽出された因子が能力を示し，かつ能力を推定するのに因子得点の方程式が適当な方法である事を示した。

1 Introduction

The abilities or traits that we wish to study are usually not directly measurable ; rather they must be studied indirectly, through measurement of other quantities, such as motor performances. Moreover, the concept "ability" has been constructed by us to explain or investigate the characteristics and nature of human and it is not visible. For an example, we cannot directly measure a person's explosive strength ; we can only measure his performance which he could achieve with his best exertion of explosive strengh. Furthermore, as far as we try to study the ability of interest through measuring the performance, it must be confirmed that the performance measured is achieved by exertion of only the ability of interest. However, most of motor performances are usually accomplished by exertion of various abilities, although their degrees of contribution to motor performance are different. Vertical jump is the one of adequate items to measure explosive strength, so it has been widely used to measure the explosive strength. In order to perform vertical jump, however, explosive strength must be exerted but other kinds of abilities, such as coordination of limbs, coordination of limbs and sight, and then, body size ; stature, uppoer limb length, and body weight also relate the performace of vertical jump, though their degree of contribution to the perormance are not equal. Vertical jump is rather simple motor performance but several kinds of abilities and body size or physique have to do with the performance of vertical jump. Therefore, it is almost impossible to choose the motor performance accomplished by only the ability of interest. This implies that the results obtained by measuring the motor performance do not show the ability of interest itself directrly.

Another thing we have to notice is that the

* 中京女子大学　大府市横根町（〒474）
Chukyo Women's University　Yokone, Ohu 474

motor performace test scores contain some sizable errors of measurement. In physical sciences we can usually repeat the same measurement a number of times, but in ability testing, such as test of physical fitness, motor ability and motor skill, we can probably repeat a measurement once or twice. If we attempt further repetitions, the testee's performance may change substantially because of fatigue or practice effects.

From two reasons mentioned previously, the theory of measuring ability is necessary in sport science. Anyway, it is common in various fields of study that we postulate the theoretical constructs that are not directly measurable and we observe the phenomena that manifest these theoretical constructs and we can make only inferences on these theoretical constructs by the scores or results obtained.

2 Definition of measurement

Measurement is a procedure for the assignment of numbers or symbols to specified properties of subjects in such a way as to characterize and preserve specified relationships in the domain of ability of interst. Then, the assignment of numbers or symbols provids that the numbers are logically and naturally amenable to analysis by mathematical operation according to certain specified rules. For assignment of numbers to the specified properties of subjects, the scales used very often in physical sciences, such as CGS-unit, are very often used in sport science, too. In other words, measurement can be defined as a procedure for the assignment of numbers or symbols to specified properties of subjects with such scale as popularly used in physical sciences. Evaluation, however, is very similar to measurement in terms of procedure, but the scale used for evaluation is the "value scale". Thus, evaluation is a procedure for the interpretation of specified properties of subjects with "value scale".

3 Model of ability

Lord and Novick described that measurement begins with a procedure for identifying the elements of the real world with the elements or constructs of an abstract logical system (a model) through the precise semantic definition of the basic elements of the theory.[3] There are usually two kinds of models in most of studies. They are mathematical model and verbal model. Then, Lord and Novick also pointed the differences between these two kinds of models ; (1) It is identified with an exact mathematical system, usually of a very high order (an algebra or calculus) , by which the elementary constructs may be manipulated to facilitate deduction from the model, (2) Since the mathematical model is more precise, its use avoids the confusion that results from the imprecise statement of the purely verbal model, (3) The constructs of system tend to have less connotative meaning and thus, perhaps, have less conceptual richness, because the constructs of system are more explicitly defined, (4) The mathematical model is usually abstracted from and portrays only very limited aspsects of behavior domain.[3] The theoretical constructs posturated are often related to the motor performances through observable variables by considering the observable variables as "measures or indicants" of the theoretical constructs ; ability. And, conversely, theoretical constructs are often abstracted from given observable variables. Then, if the expected value of observable variable can be presumed to increase monotonically with the theoretical construct, the observable variable can be defined "measure" of thoretical construct ; ability.

Let θ be an ability of interest as theoretical construct and X be an observable variable. If X is "measure" of θ , X can be expressed by a certain function ϕ of θ , although ϕ is not necessarily linear.

$$X = \phi(\theta) \qquad (1)$$

where X is a dependent (observable) variable and θ a independent (unobservable variable).

In physical sciences, the researchers generally find the θ available to them does not compleltely determine X ; i. e., for fixed θ there is still some variation among X values they observe. Whenever for the fixed θ and the specified function ϕ the remaining variation of X is small enough to be neglected, this function ϕ is "adequate" in terms of the degree of accuracy that might in practice be desirable in monitoring X. In this case, we can usually say that θ accouts for most of variation in X. This equation is the statement of "deterministic model" ; it asserts that θ determines X. This model, however, has found only limited use in sport science domain.

In practice, even if a very large number of parameters $\theta = (\theta_1, {}_2, \ldots, \theta_k)$ are used to describe X, we usally are unable to determine X very exactly. Then, we need another kind of model ; that is,

$$X = \phi(\theta) + e \qquad (2)$$

where e, called an error (or residual), is considered to be a composite of effects not associated with the available independent variables θ. This model is called "probabilistic model". In sport scientific domain, though the determination of ϕ is not so easy, this "probabilitstic model" is more applicable and reasonable to the study of measuring such ability as physical fitness, motor ability, motor skill and so on, because a large number of repetition can not be made for measuring such abilities.

4 Ability

Any theory of ability supposes that a individual's motor performance can be accounted for, to a substantial degree, by defining certain human characteristics called "ability or trait", quantitatively estimating the individual's standing on this ability, and then using the numerical values obtained to predict the individual's performance in relevant situation. Much of scientific field studying ability is based on an ability orientation, but nowhere is there any necessary implication that ability exists in any physical or physiological sense. It is sufficient that a person performs as if he were in possession of certain amount of ability of interest and he performs as if this amount of ability substantially determined his motor performance.

Now let us consider a set of n test items and a set of k kinds of ability. Then we assume that each of the abilities affects the subject's motor performance on at least one test item in the set. These abilities are to be thought of as physical fitness or motor ability domains necessary for the description of individuals. We shall denote them by the vector

$$\Theta = (\theta_1, \theta_2, \ldots\ldots, \theta_k) \qquad (3)$$

We can now represent each person by a point in k-dimensional space, called an "ability space".

The regression of test item score on Θ is called the "item characteristic function", just same as in psychometics. This item characteristic function is a key concept for making inference about unobservable ability from the observed test score. However, the item characteristic function cannot be observed directly from the simple reason that Θ is unobservable.

5 Factor extracted by factor analytic procedure as ability

Suppose that there are n test variables which are validated to measure the identical ability ; T_i $(i = 1, 2, \ldots\ldots n)$. These test variables are by nature highly correlated each other. And we postulate that these high correlations are obtained because these T_i measure same ability θ. Suppose that X_i $(i = 1, 2, \ldots, n)$ are test scores of T_i. This postulate means that $X_i = f_i(\theta) + e_i$ is held for all X_i. If we can determine these functions f_i in least square sense, these f_i are equivalent to the item characteristic functions.

Now let us discuss about more general case. Under the postulate mentioned previously, the correlation between two test variables means to which extent these test variables measure the identical ability. Let $X' = (X_1, X_2, \ldots, X_n)$ be row vector of n test variables and also $\Theta' = (\theta_1, \theta_2, \ldots, \theta_k)$ be row vector of k ability variables and usually $k < n$. We denote the dispersion (variance-covariance) matrix of X by

$$\Sigma = [\sigma_{ij}] \quad i=1,2,\ldots,n, \quad j=1, \quad 2,\ldots, \quad n$$

The dispersion matrix of Θ by

$$\Psi = [\psi_{ij}] \quad i=1,2,\ldots,k, \quad j=1, \quad 2,\ldots, \quad k$$

Then, let $\Lambda = [\lambda_{ij}]$ $(i=1,2,\ldots,n, j=1,2,\ldots,k)$ be an $(n \times k)$ matrix of constants (weights relating X to Θ) called "factor loadings". And let us require that these constants are such that each column of Λ has two or more nonzero elements. This requirement identifies each element of Θ as a common factor.

A basic equation of the factor analytic model is

$$X = \Lambda \Theta + U, \qquad (4)$$

where U means the remaining part which is not explained by Θ, called uniqueness of X in which error is involved. The equation (4) is just comparable to the item characteristic function (2). Therfore, if we can determine Λ in this equation, the item characteristic function can be determined. Then, the procedure to determine Λ involves a decomposition of Σ into two parts ;

$$\Sigma = \Lambda \Psi \Lambda' + UU'. \qquad (5)$$

That Ψ is identity matrix means that the elements of Θ are uncorrelated each other ; independent each other. In this case, this equation is expressed as follows ;

$$\Sigma = \Lambda \Lambda' + UU'. \qquad (6)$$

Λ of this kind is called "orthogonal solution" in factor analysis. This decompsition of Σ can be done by determination of eigen values and vecotrs of Σ and then Λ is determined. The relative magnitudes of the elements of any row of Λ, the vector of factor loadings, determine the relative extent to which that test variable measures each of abilities.

Then, let $\Theta = BX$ be equations for estimating abilities Θ by the observable variables. From the basic equation of factor analysis (4),

$$\Theta = B(\Lambda \Theta + U) \qquad (7)$$

Then, postmultiply $(\Lambda \Theta + U)'$ to both sides.

$$\Theta(\Lambda \Theta + U)' = B(\Lambda \Theta + U)(\Lambda \Theta + U)'$$

Mathematical expectation of both sides is

$$\varepsilon\{\Theta(\Lambda\Theta + U)'\} = B\varepsilon\{(\Lambda\Theta + U)'(\Lambda\Theta + U)'\}$$
$$\varepsilon(\Theta\Theta')\Lambda' + \varepsilon(\Theta U') = B\{(\Lambda\varepsilon(\Theta\Theta')\Lambda' + \varepsilon(U\Theta')\Lambda' + \Lambda\varepsilon(\Theta U') + \varepsilon(UU')\}$$

In this equation, $\varepsilon(\Theta U') = \varepsilon(U\Theta') = 0$ and $\varepsilon(UU') = D^2$ is the dispersion matrix of U and it is diagonal matrix whose diagonal elements are the variances of elements of U. And $\varepsilon(\Theta\Theta')$ is the dispersion matrix of Θ, so if the orthogonal solution is chosen, $\varepsilon(\Theta\Theta')$ is an identity matrix ; I. Thus, this equation is

$$\Lambda' = B(\Lambda\Lambda' + D^2) \qquad (8)$$

Therefore,

$$B = \Lambda'(\Lambda\Lambda' + D^2)^{-1} \qquad (9)$$

Thus

$$\Theta = \Lambda'\{\Lambda\Lambda' + D^2\}^{-1} X \quad (10)$$

is the equation for estimating Θ by the observable variables X.

6 Problem

The theory of ability from factor analytic model can be derived so elegantly in mathematical sense, but there is one problem which we have to keep in mind.

Suppose that (Θ, Λ, Ψ) provides a factor analytic decomposition for some value k ; number of abilities. Let T be an arbitrary nonsingular matrix of order k, and let

$$\Lambda^* = \Lambda T^{-1} \tag{7}$$
$$\Psi^* = T \Psi T' \tag{8}$$
$$\Theta^* = T \Theta \tag{9}$$

Then, $\Lambda^* \Theta^* = \Lambda \Theta$, and hence

$$X = \Lambda^* \Theta^* + U$$

and Ψ^* is the dispersion matrix of Θ^*. Therefore, (Θ^*, Λ^*, Ψ^*) provides a factor analytic decomposition of X different from that provided by (Θ, Λ, Ψ). In general if a factor analytic decomposition exists for some value of k ; number of abilities, then there must be an infinite number of solution for that k. The problem of which decompostion to choose is called the "problem of rotation". To limit the number of solutions that may be possible for any value of k to finally a unique decomposion, we may place various restrictions on the model, in the form of conditions that the ability variables Θ must satisfy. For an example, we can limit the number of solutions somewhat by requiring that the common factors Θ be "orthogonal", that Ψ be an identity matrix. The solution in which Ψ is not equal to an identity matrix is called the "oblique solution", as well known, but there will always exist a transformation that yields an orthogonal factorization when used with equations (7), (8), and (9), mentioned previously. The theory of simple structure of factor loading matrix is another restriction to identify a unique decomposition.[1)5)] Actually, however, the final determination of decomposition is subject to the researcher's subjective judgement.

References

1. Harman HH (1962) : Modern factor analysis, The university of Chicago press, Chicago, pp. 112-113.
2. Lazarsfeld PF (1950) : The logical and mathematical foundation of latent structure analysis, in Measurement and prediction, ed. by Stouffer SA, Princeton University press, New Jersey, pp. 362-412.
3. Loard FM and Novick MR (1968) : Statistical theories of mental test scores, Addison-Wesley Publishing Company, Massachusetts, pp. 10-12.
4. op. cit., 19
5. Matsuura Y (1972) : Factor analysis in behavioral sciences, Fumaido, Tokyo, pp. 216-217.
6. Matsuura T (1980) : Multivariate assessment of physical fitness, in Kinanthropometry II , ed. by Ostyn M, Beunen G and Simons J, University park press, Baltimore, pp. 161-175.