Journal of Human Kinetics volume 10, 2003, 99-106

The Rasch model in the motor ability testing

by
*Ladislav Cepicka**

Motor abilities are essential for the execution of many motor tasks. The purpose of this study was to introduce the Rasch analysis in the context of motor ability testing. To investigate the application of the Rasch model for the evaluation of motor abilities the Brace test (BT) was chosen. The subjects included 105 university students aged from 19 to 24 years. Test performance for all students on the BT was subjected to analysis for fit to a theoretical Rasch measurement model. The results indicate that BT is relatively homogenous with respect to its content. In addition Rasch analysis in the evaluation of motor abilities was examined. According to our results we can conclude that the Rasch model is a very powerful instrument in motor ability assessment.

Keywords: *Rasch model, Brace test, motor abilities*

* University of West Bohemia, Department of Physical Education and Sport Science, Klatovská 51, 301 00 Plzen, The Czech Republic

Introduction

The question of a motor ability investigation represents a difficult problem because of the large variability of variables. Humans are objects of strong individual traits. Moreover, many of those traits are difficult to measure. There is a need to investigate them to get a feedback in the domain of physical education and sport. The knowledge of abilities structure helps to select better methods of its development.

Motor tests are used to evaluate the level of particular abilities. There are two main theories to quantify motor abilities: (1) Theory well known as a classical test theory. This theoretical approach has been usually applied to norm-referenced measurement problems and it is based on the normal distribution principle. (2) Recently, criterion-referenced methods have become popular (especially in physical education) although only a few published tests (on this approach) have been available. They have been accepted by physical education teachers, especially those espousing a mastery learning instructional model (Safrit 1987). A few theories exist in this frame, but one of them has been focused more than the others. It is the item response theory (IRT) also known as latent trait theory. IRT has been developed to obtain more information from the tests results. It offers more advantages than traditional procedures. The main advantage is to obtain an estimate of an unknown ability and to estimate the influence of one independent variable on another. Another advantage is the same scale for the measurement of the item difficulty and ability. A plot of these variables is called the item characteristic curve. Usually, we can describe one, two or three parameters: a discrimination parameter, a difficulty parameter and a guessing parameter. There is one model, which uses the assumption that for the description of the characteristic curve just the difficulty parameter is important. This, one-parameter model is called a Rasch model. The Rasch model is very popular and widely used in psychomotor evaluation, appropriate for dichotomous data. This model appears to be useful for tests of motor behavior consisting of several trials, which are scored on 0-1 basis (McDonald 1985). There are not too many applications in the domain of motor behavior. Thus, the purpose of this study is to investigate the possibilities of the Rasch model for the measurement of skills and abilities in physical education.

This paper presents the application of the Rasch model. This one is the latent trait model proposed for binary data analysis and we can describe it by nonlinear functional relationship, which is expressed by the equation:

$$P_j(F) = \frac{e^{(f_j - b_j)}}{1 + e^{(f_j - b_j)}} \qquad (1)$$

The elements of this equation are called person and item parameters. $P_j(F)$ is the probability of correct response for item *i* at ability *F*. The parameter *b* is called the difficulty index or just item difficulty and that is point on the ability scale at which the probability of correct response is 50 percent. Item parameters describe item characteristic curve that indicates the form of the relationship between the observable response to the item and unobservable trait or ability level of an individual. This curve is illustrated in fig. 1.

The latent ability is represented on the horizontal axis and probability of correct response on the vertical axis. This curve shows probability $P_j(F)$ for any level of ability *F*. In the case of the Rasch model the curves of every items have the same shape but they are shifted more below or up to the mean level of ability. The curve in fig. 1 is an example when the level of 0 ability corresponds to 50 percent probability of correct response. Thus item characteristic curve is the nonlinear regression of the item score at each ability level on the measured ability. Although the ability scale extends from $-\infty$ to $+\infty$, the range of commonly obtained values is -3 to +3.

According to Spray (1987) there are two basic assumptions of the Rasch model:

1. The unidimensionality assumption. It is assumed that the latent trait (or ability) represents just one dimension.
2. The local independence assumption. It is assumed that the responses to two items are unrelated for a given level ability *F*. It supposes non-correlated errors across item and persons for the same ability level. Thus, Rasch model gives us an opportunity to estimate the subject ability and the item difficulty independently of each other.

The question of precision with which the given ability is estimated is quantified by standard error of estimate of ability. Generally, it is possible to state, that the smaller standard error of estimate, the more precise the estimate of ability.

Material and methods

The subjects included 105 students of University of West Bohemia aged from 19 to 24 years. They have been chosen at physical activity classes and courses. According to Safrit (1987) the minimum sample size needed to use a one ? parameter model for dichotomous data is 100 ? 200.

The Štepnicka's modification of Brace test (BT) was chosen to investigate the use of Rasch model for the measurement of motor ability. The Brace test corresponds to the rate of motor learning and the level of coordination abilities. It originally consisted of 21 exercises. Each exercise represents one motor task (item) of different difficulty. In this study, we have used Stepnicka's modification, which consists of ten exercises (Štepnicka 1976). Each task is presented and explained before test protocol. The best result was registered. Task is scored by 2 ? successful at first attempt, 1 ? successful at the second attempt, 0 ? unsuccessful. This study focused only on motor control so the test was scored: 1 ? successful, 0 ? unsuccessful. The test score is the sum of the item scores.

Three examiners, all members of Department of Physical Education and Sport Science, were trained in the administration of the test. To record the test score a special form was prepared. The test protocol has been discussed before testing.

The Germano Rossi's Rasch 1.0© software was used to apply the Rasch model and to calculate all item difficulty and person ability estimates as well (Rossi 1991). These estimates were evaluated in two ways: (1) The quality of fit of the observed data to the model was calculated using χ^2 technique. This one is based on answers of all subjects on the item. (2) The standard errors of estimate were calculated for all values.

Descriptive statistics were calculated for difficulty and ability estimates and test scores. Mean, standard deviation, maximum and minimum ability values as well as standard error of estimate were obtained. Item and test information function were used to examine the precision of the measurement of specific levels of ability.

Results

Overall the model fits the data very well. The Brace test had just 6 misfit persons of 105 subjects (tab. 1-2). The standard error of estimate (SEE) for the ability and difficulty estimates were generally small. The mean SEE of the ability estimates was 0,853. For the difficulty estimates, the mean SEE was 0,242 and fell to the range from 0,220 to 0,337. The item difficulty parameter was estimated with greater precision than the ability parameter. The descriptive statistics are in the table below. The maximum ability was 2,563 and minimum was -1,617, the raw score was 5,53.

Table 1 The descriptive characteristics

T	Ability	Difficulty
Mean	1,248	0,00
Max	2,563	2,474
Min	-1,617	-1,190
SEE mean	0,583	0,242
Raw score mean	5,53	N/A

Table 2 The item parameters

Item	Si	difficulty	SEE	c²	d.f.	p
1	81	-1,071	0,257	46,432	95	1,000
2	83	-1,190	0,264	79,865	95	0,867
3	48	0,457	0,221	79,865	95	0,880
4	57	0,073	0,221	62,074	95	0,996
5	72	-0,599	0,235	70,628	95	0,971
6	64	-0,231	0,224	67,163	95	0,986
7	63	-0,186	0,224	57,816	95	0,999
8	55	0,158	0,220	55,043	95	0,999
9	56	0,116	0,220	57,382	95	0,999
10	13	2,474	0,337	73,714	95	0,948

where Si means number of 1 score on the item, d.f. means degrees of fredom.

Table 3 The probability values

r	1	2	3	4	5	6	7	8	9	10	ability
2	0,37	0,39	,011	0,16	0,27	0,20	0,19	0,14	0,15	0,02	-1,62
3	0,52	0,55	,019	0,26	0,40	0,32	0,31	0,24	0,25	0,03	-0,99
4	0,65	,067	0,28	0,37	0,53	0,44	0,43	0,35	0,36	0,05	-0,47
5	0,74	0,77	0,39	0,48	0,65	0,56	0,55	0,46	0,47	0,08	0,00
6	0,82	0,84	0,50	0,60	0,74	0,67	0,66	0,58	0,59	0,12	0,47
7	0,89	0,90	0,63	0,71	0,83	0,77	0,76	0,70	0,71	0,18	0,99
8	0,94	0,94	0,76	0,82	0,90	0,86	0,86	0,81	0,82	0,30	1,62
9	0,97	0,98	0,89	0,92	0,96	0,94	0,94	0,92	0,92	0,52	2,56
Diff.	-1,07	-1,19	0,46	0,07	-0,60	-0,23	-0,19	0,16	0,12	2,47	

The item difficulty parameter was within a range from -1,071 to 2,474. The easiest tasks were 2 and 1 which 83/81 subjects scored successfully. The common feature of those is some kind of balance when the center of gravity is near the floor and arms are spread out. The item 4 cross leg squad had almost average difficulty and the probability of successful score was close to 50%. The most difficult task was 10. This is a really tough exercise for which coordination, flexibility, strength and courage are needed. The difficulty of item 10 was 2,474.

The probability values on each item are presented in the table below. They are placed in the columns and represent the probability of success on the task. The values in the rows describe the probability of success on item for the number successful item *r*.

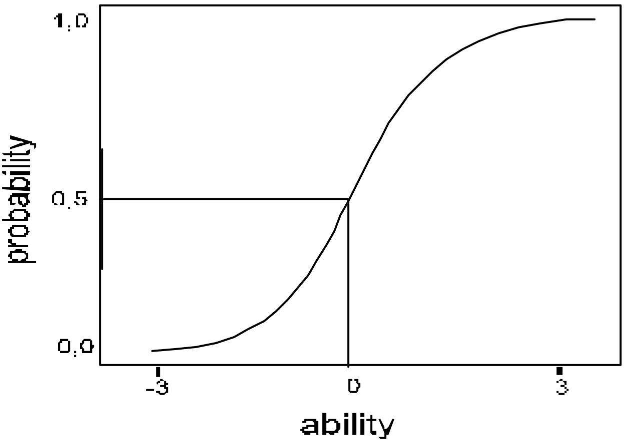

Fig. 1 Item characteristic curve

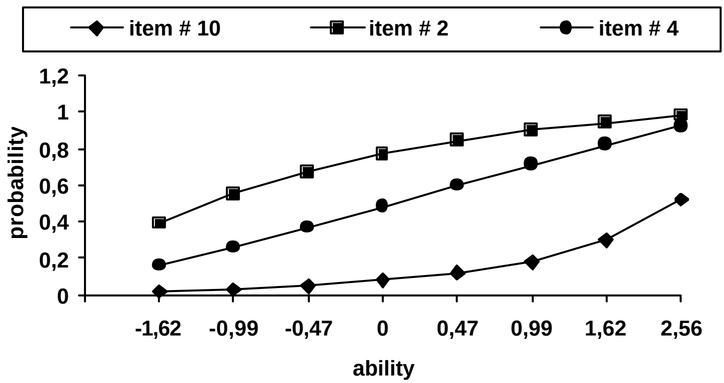

Fig. 2 Item characteristic curves

The characteristic curves of three items are illustrated in Fig. 2. The most difficult item is number 10 and the easiest number 2. The characteristic curves of the other items should be in the area between both curves. As we can see, the difficulty of item 10 is high. The subject with the average level of ability has just close to 10% chance of success. On the other hand the same person will be successful on item 2 with the probability close to 75%. The difficulty parameter of item number 4 is close to 0 ability and the probability is close to 50%.

Conclusion

The aim of this study was to examine empirically the usefulness of the Rasch model for a specific type of item in psychomotor domain. The conclusions were that the model-data fit was generally good across persons and across trials. When the model-data fit is good, the Rasch model can be used for parameter estimation (Safrit, 1987). The item parameters were estimated more precisely than the person ability parameters. The majority of subjects responded in ways predicted by the assumptions of the Rasch model. The misfit persons were not removed from this sample because the study was exploratory. The number of misfit person's values was small because the Brace test had been standardized before.

The item difficulty values set have an important contribution. It gives us an opportunity to change the order of items in the test from easiest to most difficult. It can be important for some kind of „warm-up“ effect or for motor learning. Some tasks are similar and transfer can occur. It is possible to construct some kind of Guttman's scale on this base.

The Rasch model seems to be a very good method for the estimation of person ability and item (or task) difficulty in the domain of motor tests with dichotomous scores.

References

Disch J. 1987. Recent Developements in Measurement and Possible Applications to the Measurement of Psychomotor Behavior:A Response.Research Quarterly for Exercise and Sport, 58, 210-212.

Fischer G.H. 1995. Some neglected problems in IRT. Psychometrika, 60, 459-487.

McDonald R.P. 1985. Factor Analysis and Related Methods. Lawrence Erlbaum Associates.

Rossi G. 1991. Rasch 1.0 (c). Software guide.

Safrit M.J. 1987. The Applicability of Item Response Theory to Tests of Motor Behavior. Research Quarterly for Exercise and Sport, 58, 213-215.

Safrit M.J., Cohen A.S. and Costa M.C. 1989. Item Response Theory and the Measurement of Motor Behavior. Research Quarterly for Exercise and Sport, 60, 325-335.

Spray J.A. 1987. Recent Developements in Measurement and Possible Applications to the Measurement of Psychomotor Behavior. Research Quarterly for Exercise and Sport, 58, 203-209.

Stepnicka et al. 1976. Somatotype, body posture, motor level and motor activity of youth. Acta universitatis carolinae gymnica, 12, 11-79.

Acknowledgment:

This research was supported by a grant from the Grant Agency of the Czech Republic no. 406/02/P133 awarded to the author.