ACTA UNIVERSITATIS CAROLINAE
KINANTHROPOLOGICA Vol. 47, 1 – 2011 Pag. 5–15

CHARLES UNIVERSITY IN PRAGUE,
FACULTY OF PHYSICAL EDUCATION AND SPORT,
DEPARTMENT OF KINANTHROPOLOGY[1]
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN,
DEPARTMENT OF KINESIOLOGY AND COMMUNITY HEALTH[2]

INFLUENCE OF CATEGORIZATION ON INTERNAL STRUCTURE OF AN EXERCISE BARRIER INVENTORY

JAN ŠTOCHL[1] & WEIMO ZHU[2]

SUMMARY

Based on a collapsing procedure and Rasch analysis, a Rasch-based optimal categorization procedure has been introduced for the determination of the categorization of a test or scale. As a result of the Rasch-based optimal categorization, the number of response categories could be reduced, which may be a threat to the internal structure of a measure. Using three available data sets (two data sets from one study with $N = 480$ and third set from another study with $N = 219$), this study examined the differences among the structures of exercise barrier scale constructs when a different number of response categories is used. Specifically, two models of exercise barrier constructs were compared using the structural equation modeling. The results suggest that the collapsing of categories has no effect on the structure of the latent variables. In addition, the results suggest that the collapsed number of categories provides a slightly better model-data fit statistics. Two consequences for the no-impact finding are: (a) a better categorization may help eliminate systematic error related to response categories and (b) the range of ability, or between-subject variance, was still maintained. More studies are needed to determine these possible explanations' contributions. The analysis of the internal structure illustrated in this study should be a part of the Rasch-based optimal categorization procedure.

Key words: Exercise Barrier Inventory, categorisation, validity, factor analysis, item response theory

INTRODUCTION

Like a brick in a wall, an item is the basic unit of any test or scale. The item could be a statement, command or question that states is associated with the desired trait/ability to be measured in clear unequivocal terms. An item often contains a series of responses, which is known as "selected-response," or it may require a respondent to construct his/her own responses, which is known as "constructed-response." While multiple-choice, true-false, short-answer, and essay are commonly used item formats in educational measurement

practice, the rating scale format is the most popular one in psychological measurement practice. A rating scale usually includes a statement and a set of response categories, e.g.:

Statement: *Time is never an issue for my regular exercise.*

Response options/categories: *Strongly Disagree, Disagree, Agree & Strongly Agree*.

Because of its role in a test or scale, development of items is considered one of the most important aspects in test or scale construction, and a number of books (e.g., Haladyna, 1994; Osterlind, 1989; Roid & Haladyna, 1982) have been devoted specifically on how to develop good test items. In fact, item development is believed to be a part of construct validation (Haladyna, 1994), and Thorndike (1967) even called the test item a type of "mini" test. Wiggins and Goldberg (1965) proposed a new field of study called "itemmetrics" to study measurement properties of individual items. Specifically, they believe that item properties, such as ambiguity, social desirability, grammatical form, retest consistency, etc., should be examined for each inventory being developed. After a review of research on item writing, item format, test instructions and item readability, Benson (1981) concluded that careful consideration in item development will lead to the enhancement of the content validity of test score interpretation and thus to improvements in measurement of student performance and in evaluation of educational programs.

For a rating scale item, categorization is a very important part in the item development. A number of factors, e.g., number of categories (Miller, 1956), label and position (Klockars & Yamagishi, 1988), and type of anchor (Wedell, Parducci, & Lane, 1990), may have a potential impact on the categorization. For the categorization in a rating scale, there are two important features: (a) each category must has its own well defined boundaries although all categories are measuring the same trait or ability and (b) categories must be in an order and numerical values generated from the categories must reflect the degrees or magnitudes of the trait (Zhu, Updyke, & Lewandowski, 1997). An optimal categorization is a one that best exhibits these characteristics.

Traditionally, quality of an item is determined using conventional item analyses, e.g., item difficulty, discrimination, and internal consistency by Cronbach's alpha, but they provide limited information on appropriateness of categorizations. Fortunately, by combining a collapsing procedure and Rasch analysis (Wright & Masters, 1982), a new optimal categorization procedure was proposed (Zhu et al., 1997), which includes four steps:

1. combine adjacent categories in a "collapsing" process, in which new categorizations are constructed;
2. select an appropriate Rasch model, apply the Rasch calibrations, and examine the model-data fit;
3. if the model-data fit is satisfactory, identify the "candidates" of the optimal categorization whose categories are ordered;
4. determine the optimal categorization by selecting it from the "candidate" categorizations exhibiting the greatest statistical separation.

The procedure has proven to be a useful means in determining the optimal categorization of an ordered-response scale (Zhu, Timm, & Ainsworth, 2001; Zhu, et al., 1997), shown its stability when the scale was applied to a cross-cultural sample (Zhu & Kang, 1998) and been confirmed in a follow-up study (Zhu, 2002).

The impact of optimal categorization on the internal factorial structure of a test or scale, however, has not been studied. Because categorization is an essential element of an item,

which in turn, is the basic units of a test/scale, the completion of the optimal categorization should lead to a clearer factor structure of a test or scale. Although this deduction seems reasonable, it had yet to be examined empirically. This study addressed this need with an examination of the influence of the optimal categorization on the factor structure of an exercise barrier inventory.

METHODS

Exercise Barrier Inventory and Data

Two data sets (N = 480 each) of an exercise barrier inventory from a previous study by Zhu et al. (2001), one set of 5-point original data and another set of 3-point data collapsed from the original one, were used for examining the influence of optimal categorization on the factor structure. The inventory consists of 23 items scored originally on a 5-point Likert-type scale (1 = "Very often", 2 = "Often", 3 = "Sometimes", 4 = "Rarely", and 5 = "Never"). According to a principle component analysis (Zhu, 2002), these barrier items can be summarized into six domain categories, or factors, including "Resources/Skills", "Psychosocial", "Personal Well-being", "Time", "Weather/Inconvenience", and "Family/Friend Support" (see Table 1). The results of the Rasch-based optimal categorization study (Zhu, et al., 2001) suggested collapsing response categories into 3-point scale: 1 = "Very often/Often", 2 = "Sometime/Rarely", and 3 = "Never". This 3-point categorization was confirmed by a follow-up study by Zhu (2002), in which the inventory with the new categorization (1 = "Very often", 2 = "Sometimes", and 3 = "Never") was administered to a subsample (n = 219) of the original study sample and the item and category statistics were compared to those from the original inventory with the 5-point categorization. The availability of the original 5-point and collapsed 3-point data made it possible to study the influence of the optimal categorization on factor structure (see also Table 1). Note that for the convenience of the model testing, Item "inconvenience of perspiration or combing" was slightly modified for its domain in the below factor analyses: as the one which does not belong to weather but rather a standalone item.

Table 1. Items and domains of exercise barrier inventory (Guan & Zhu, 1999)

Domain	Item or factor name	Abbreviation
Resources or skills		resource
	Cost of exercising	cost
	No facilities or space to exercise	faciliti
	Lack of equipment	equipmen
	Not knowing how to exercise	exercise
	Lack of skill	skills
	Lack of transportation to get to place to exercise	transpor

Domain	Item or factor name	Abbreviation
Psychosocial		psy-soc
	Lack of interest in exercise or physical activity	interest
	Lack of enjoyment from exercise or physical activity	enjoymen
	Lack of self-discipline	discipli
	Discouragement	discoura
	Self-conscious or embarrassed about my looks when I exercise or do physical activity	consciou
Personal well being		well
	Bad health	health
	Pain or discomfort	discomfor
	Fear of injury	injury
	Exercise intensity required to improve health is too high for me	intensit
	Fear of safety	safety
Time		time
	Lack of time	notime
	Lack of block of time for doing exercise	bltime
Weather or inconvenience		weather
	Bad weather	badweath
	When is too cold or too hot	cold
	Inconvenience of perspiration or combing	inconven
Family or friend support		family
	Lack of company	company
	Lack of family support	familysu

Structure Equation Modeling (SEM)

Two SEM models were employed under the scope of this study. The first one was based on the items with 5-category responses and the second included the same items, but with collapsed response categories (3-categories) recoded in the same way as described above. Both models were tested separately to obtain fit indexes. Subsequently, the multigroup approach was employed to compare the models simultaneously. The six-factor structure of the model followed the study by Zhu (2002), i.e., "Resources/Skills", "Psychosocial", "Personal Wellbeing", "Time", "Weather/Inconvenience", and "Family/Friends support". In addition, a higher order unidimensionality has been proposed. The factor has been denoted as "Barriers". The structure is illustrated in Figure 1.

For the purposes of this study, the LISREL program (Jöreskog & Sörbom, 2005) was used. Since the data is on the Likert scale and the sample size is relatively small, Jöreskog

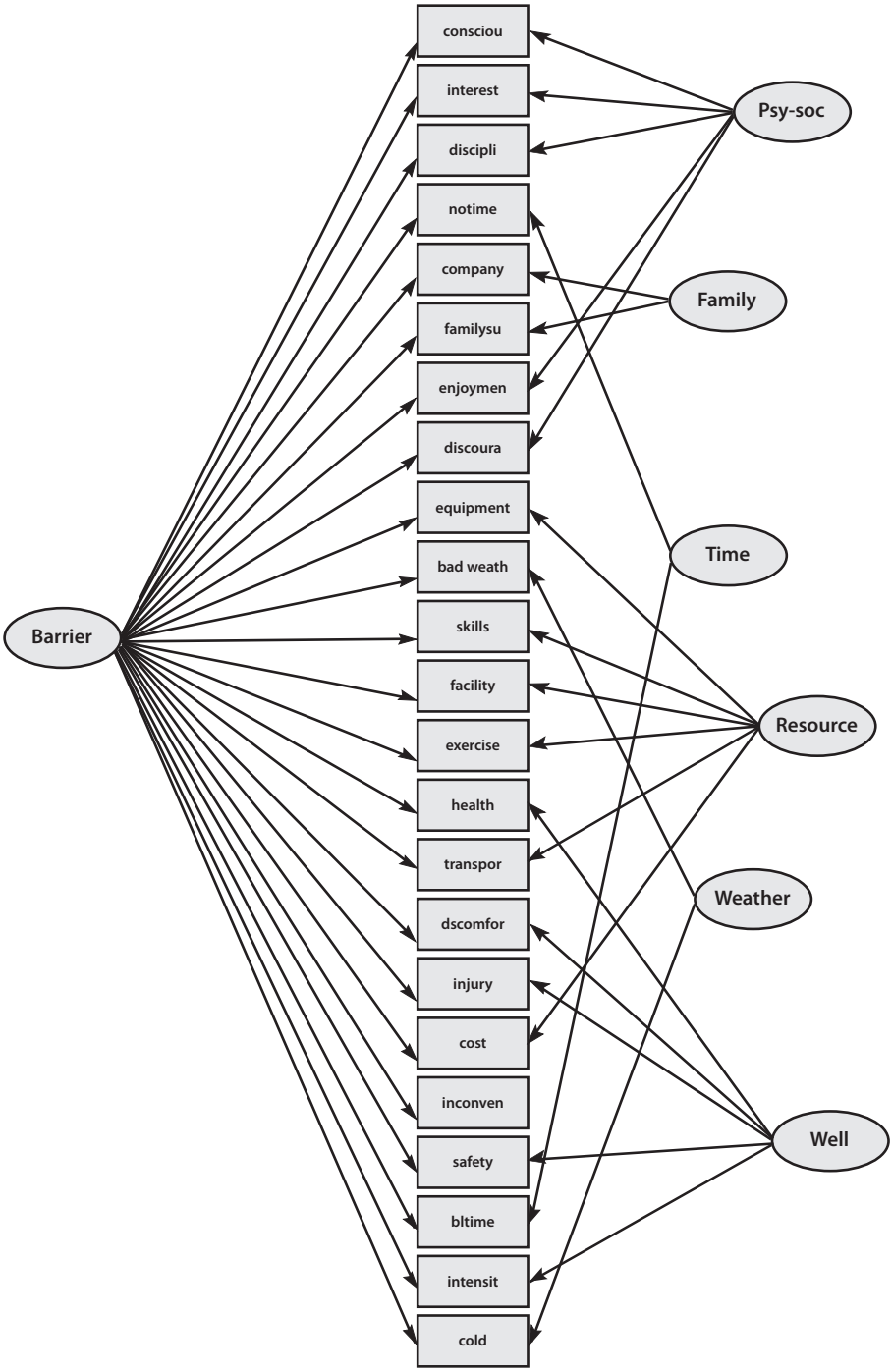

Figure 1. Path diagram of the 6-factor model of exercise barrier scale

& Sörbom (1993) recommend to analyze the matrix of estimated polychoric correlations of the observed variables together with corresponding matrix of asymptotic covariances, and to use the robust Diagonally Weighted Least Squares (DWLS) method for parameter estimation. Polychoric correlations and asymptotic covariance matrix of those estimated correlations were computed using the PRELIS program (Jöreskog & Sörbom, 2002).

As recommended by Boomsma (2000), the matrix of estimated polychoric correlations, substantive goodness-of-fit indices, such as the root mean square error of approximation (RMSEA), Satorra-Bentler's chi-square statistic, standardized root mean square residual (SRMR), Akaike's information criterion (AIC), normed fit index (NFI), comparative fit index (CFI), goodness of fit index (GFI), and summary of estimated standard errors of the parameter estimates and residuals, are reported. For the multi-group analysis the chi-square test of difference is presented.

RESULTS

Correlation Matrices

Table 2 shows the impact of categorization on the polychoric correlations. Correlations of 5-point categorization are above the diagonal, 3-point categorization correlations are below the diagonal. Obviously the impact of categorization on the values is minimal.

Fit of the Models

Table 3 provides fit indices of both models (5-categories and 3 categories). In general, both of the models fit the data well. Values of RMSEA, model AIC, Sattora-Bentler Chi-square, and CFI suggest the slightly better fit of the 3-point scale model. On the other hand, the range of unexplained correlations (fitted residuals) and SRMR is slightly better for the 5-point scale model.

Table 3. Fit indices of 5-point and 3-point scale

Fit statistics	5-point scale	3-point scale
Sattora-Bentler Chi-square	315.20	313.47
df (p-value)	208 (P = .00)	208 (P = .00)
RMSEA	.033	.033
GFI	.98	.98
Model AIC	451.20	449.47
CFI	.99	.99
NFI	.98	.98
SRMR	.060	.063
Fitted residuals range	(–.17, .25)	(–.21, .29)
Standard errors range	(.03, .13)	(.03, .12)

Table 2. Polychoric correlations of exercise barrier scale

	consciou	interest	discipli	notime	company	familysu	enjoymen	discoura	equipmen	badweath	skills	faciliti	exercise	health	transpor	dscomfor	injury	cost	inconven	safety	bltime	intensit	cold
consciou		.36	.39	.17	.39	.29	.31	.43	.39	.26	.39	.42	.33	.13	.30	.31	.32	.37	.32	.29	.15	.37	.25
interest	.38		.66	.18	.38	.26	.64	.49	.27	.26	.33	.19	.31	.12	.14	.25	.26	.15	.24	.21	.25	.43	.34
discipli	.43	.69		.32	.45	.35	.57	.48	.34	.29	.27	.27	.33	.04	.06	.15	.21	.16	.28	.16	.37	.35	.37
notime	.17	.22	.33		.24	.17	.27	.16	.20	.26	.22	.18	.13	.03	.04	.05	.07	.06	.24	.06	.76	.25	.20
company	.43	.38	.48	.26		.65	.39	.48	.43	.41	.45	.36	.41	.17	.29	.26	.30	.40	.31	.20	.28	.37	.30
familysu	.31	.27	.38	.20	.69		.38	.45	.40	.34	.38	.33	.39	.26	.40	.29	.30	.37	.29	.25	.24	.39	.28
enjoymen	.32	.66	.63	.28	.41	.40		.47	.24	.28	.38	.23	.38	.12	.13	.26	.33	.25	.32	.24	.29	.45	.37
discoura	.43	.48	.51	.17	.49	.47	.52		.34	.36	.36	.25	.37	.22	.23	.35	.37	.28	.29	.24	.27	.39	.35
equipmen	.41	.29	.37	.23	.47	.43	.27	.34		.41	.56	.65	.57	.22	.43	.27	.30	.68	.33	.37	.24	.36	.31
badweath	.30	.27	.26	.32	.43	.37	.29	.42	.47		.41	.42	.34	.24	.20	.28	.22	.32	.33	.33	.29	.28	.70
skills	.41	.36	.31	.30	.50	.43	.41	.40	.61	.51		.59	.77	.28	.49	.30	.44	.55	.37	.41	.26	.45	.36
faciliti	.45	.22	.30	.19	.40	.36	.26	.28	.67	.45	.65		.58	.28	.56	.27	.37	.58	.30	.41	.28	.38	.37
exercise	.32	.33	.35	.15	.43	.45	.41	.42	.61	.41	.79	.63		.28	.47	.31	.52	.63	.35	.39	.20	.47	.39
health	.14	.14	-.02	.03	.20	.26	.08	.22	.21	.23	.29	.32	.30		.42	.67	.50	.29	.16	.34	.11	.43	.23
transpor	.31	.17	.06	.05	.30	.41	.16	.24	.44	.22	.55	.59	.47	.48		.29	.30	.57	.22	.35	.15	.34	.29
dscomfor	.34	.26	.16	.11	.28	.31	.25	.35	.28	.27	.35	.30	.33	.69	.30		.65	.36	.24	.38	.11	.43	.37
injury	.37	.27	.22	.05	.31	.33	.31	.39	.31	.23	.42	.41	.51	.58	.35	.67		.40	.32	.59	.09	.44	.33
cost	.38	.18	.18	.06	.41	.40	.26	.27	.68	.29	.58	.59	.64	.29	.59	.36	.42		.36	.44	.19	.40	.31
inconven	.35	.24	.31	.25	.35	.34	.36	.30	.37	.40	.46	.35	.40	.22	.28	.29	.38	.41		.41	.30	.38	.36
safety	.35	.27	.21	.11	.24	.27	.24	.26	.39	.32	.49	.46	.42	.33	.38	.37	.62	.51	.50		.17	.30	.39
bltime	.17	.29	.36	.77	.32	.28	.31	.30	.27	.29	.34	.33	.23	.16	.22	.15	.10	.20	.34	.21		.32	.29
intensit	.38	.43	.36	.30	.39	.38	.45	.40	.41	.33	.46	.37	.47	.40	.34	.44	.49	.42	.40	.34	.36		.38
cold	.29	.37	.40	.24	.30	.31	.38	.37	.35	.70	.42	.39	.47	.24	.31	.39	.35	.32	.39	.43	.28	.45	

Below diagonal – collapsed (3 categories); above diagonal – non collapsed (5 categories)

Chi-square Test of Difference

Chi-square test of difference, the commonly used test for level of invariance of the model in multi-group SEM did not reject the hypothesis about the invariance of model parameters across the 5-point and 3-point scales (chi-square difference = 60.00, $df = 68$, p-value = .744). This suggests that the categorization has not affected either items' validities or error variances in the model. This finding supports the idea that the categorization did not influence the underlying structure of the exercise barrier scale.

Cross-validation of the Factor Structure of the 3-point Scale Inventory

For cross-validation purposes the new 3-point scale inventory was administered to a subsample (n = 219) of the original study (Zhu et al., 2001) and the factor structure was evaluated. Table 4 provides results of this cross-validation. Although the fit is slightly worse than for the original dataset, the fit indices suggest still a very acceptable fit of the model.

Table 4. Fit indices of cross-validation 3-point scale model

Sattora-Bentler Chi-square	491.85
df (p-value)	208 (P = .00)
RMSEA	.053
GFI	.97
Model AIC	449.47
CFI	.98
NFI	.97
SRMR	.086
Fitted residuals range	(–.26, .23)
Standard errors range	(.03, .09)

DISCUSSION

A number of efforts have been made recently to examine the scale categorization using more sophisticated statistical models and methods, such as the Rasch-based optimal categorization method (Zhu et al., 1997). In theory, after the optimal categorization is determined, the relationship among test items and test components should conform better to the internal structure; therefore, provides stronger validity evidence based on the internal structure of the measure. However, during the process of the optimal categorization, because of its post-hoc feature (i.e., the collapsing procedure did not happen till the categorization has been determined and data collected), the number of the category may be decreased. The decrease in response categories is generally believed to be a loss of information. For example, the effect of the number of response categories on the magnitude

of the Pearson correlation coefficient is well known and, in general, this coefficient tends to be lower as the number of response categories decreases (Cohen, 1983; Martin, 1973, 1978). Other studies also showed that a decrease in categories may affect the internal consistency (Bandalos & Enders, 1996; Masters, 1974), test-retest reliability and validity of the items (Preston & Colman, 2000), as well as lower interrater reliability (Cicchetti, Shoinralter, & Tyrer, 1985). Thus, to determine the impact of the categorization on the internal structure of a measure is definitely needed.

In this paper, the influence of the Rasch optimal categorization on the structure of an exercise barrier scale was studied. LISREL models fit indexes for 5-point and 3-point scales have been presented as well as simultaneous comparison (invariance) of both models. In addition, the 3-point model has been examined and compared with the data collected based on the collapsed, new 3-point scale. The results indicated the category reduction caused by the optimal categorization has not brought about any negative impact on the internal structure of the scale. In fact, a slight better model-data fit statistics have been observed. There are two possible explanations for this "no-negative-impact" observation.

First, a decrease in the number of categories is not always necessary to lead to a low internal consistency statistic. The study by Chang (1994), for example, found that more scale categories may not necessarily enhance reliability and may even lead to a systematic "abuse" of the scale. This is because some respondents may systematically skip certain responses associated with the higher number categories scales, or they may use interchangeably *very often* with *often* throughout the instrument. According to Chang (1994), both these response behaviors could contribute to the systematic error. With a better categorization, these behaviors may be eliminated. As a result, the systematic error may also be eliminated.

Second, category reduction is not to simply reduce the number of categories in those earlier category-reduction studies (e.g., Miller, 1956). Rather, the reduction, or more actually collapsing and final optimal categorization, is determined based on a set of well studied category statistics (e.g., average measure, Linacre, 2002 and Andrich's threshold, Andrich, 1978) and separation statistics (e.g., item and person separations, Wright & Masters, 1982). In other words, the reduction, if there is one during the Rasch optimal categorization, is evidence based or driven. More importantly, key factors that may impact correlation (e.g., range of the ability or between-subject variance) may not be affected at all. As an example, in a typical category reduction study, "extreme" categories (e.g., Very often and Never) at each end of the categorization will likely be deleted. This is, however, not necessary in the case the Rasch-based optimal categorization. When the number of categories is reduced from five (Very often, Often, Sometimes, Rarely, & Never) to three categories (Very often, Sometimes, Never) based on the Rasch-based optimal categorization, for example, the "extreme" categories (i.e., Very often and Never) are retained. As a result, the range of the ability being measured was also retained. Simply expecting the range being measured will be reduced because of reduced response categories is clearly incorrect. Because both factors, i.e., a better categorization and kept measurement range, are confounded in this study, more studies are needed to determine the degree of their contributions.

CONCLUSION

Although the number of response categories was reduced from a five-point scale to a 3-point scale during the Rasch-based optimal categorization procedure, the reduction did not bring about any negative impact to the internal structure of the exercise barrier instrument studied. Two explanations for the no-impact finding are: (a) a better categorization may help eliminate systematic error related to response categories and (b) the range of ability, or between-subject variance, was still maintained. More studies are needed to determine the explanations' individual contribution. The analysis of the internal structure illustrated in this study should be part of the Rasch-based optimal categorization.

ACKNOWLEDGEMENT

This study was supported by a grant from the Ministry of Education, Youth and Sports' grant, the Czech Republic (nr. MSM 0021620864).

REFERENCES

ANDRICH, D. (1978). A rating formulation of ordered response categories. *Psychometrika*, 43, 561–573.

BANDALOS, D. L., & ENDERS, C. K. (1996). The Effects of Nonnormality and Number of Response Categories on Reliability. *Applied Measurement in Education*, 9(2), 151–160.

BENSON, J. (1981). A redefinition of content validity. *Educational and Psychological Measurement*, 41(3), 793–802.

BOOMSMA, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7(3), 461–483.

CICCHETTI, D. V., SHOINRALTER, D., & TYRER, P. J. (1985). The Effect of Number of Rating Scale Categories on Levels of Interrater Reliability: A Monte Carlo Investigation. *Applied Psychological Measurement*, 9(1), 31–36.

COHEN, J. (1983). The Cost of Dichotomization. *Applied Psychological Measurement*, 7, 249–253.

GUAN, J., & Zhu, W. (1999). Validity and reliability of an exercise/physical activity barrier instrument: A preliminary analysis. *Research quarterly for exercise and sport*, 70(Suppl.), 60–61.

HALADYNA, T. M. (1994). *Developing and Validating Multiple-Choice Questions*. Hillsdale, NJ: Erlbaum.

CHANG, L. (1994). A Psychometric Evaluation of 4-Point and 6-Point Likert-Type Scales in Relation to Reliability and Validity. *Applied Psychological Measurement*, 18(3), 205–215.

JÖRESKOG, K. G., & SÖRBOM, D. (1993). LISREL 8: *Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.

JÖRESKOG, K. G., & SÖRBOM, D. (2002). *PRELIS* (Version 2.54). Lincolnwood, IL: Scientific Software International.

JÖRESKOG, K. G., & SÖRBOM, D. (2005). *LISREL* (Version 8.72). Lincolnwood, IL: Scientific Software International.

KLOCKARS, A. J., & YAMAGISHI, M. (1988). The Influence of Labels and Positions in Rating Scales. *Journal of Educational Measurement*, 25(2), 85–96.

LINACRE, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85–106.

MARTIN, W. S. (1973). The Effects of Scaling on the correlation coefficient: A Test of Validity. *Journal of Marketing Research*, 10, 316–318.

MARTIN, W. S. (1978). Effects of Scaling on Correlation Coefficients: Additional Considerations. *Journal of Marketing Research*, 15, 304–308.

MASTERS, J. R. (1974). The Relationship between Number of Response Categories and Reliability of Likert-Type Questionnaires. *Journal of Educational Measurement*, 11(1), 49–53.
MILLER, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63, 81–97.
OSTERLIND, S. J. (1989). *Constructing test items*. Boston: Kluwer Academic Publishers.
Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1–15.
ROID, G. H., & HALADYNA, T. M. (1982). *Technology For Test-Item Writing*. New York: Academic press.
THORNDIKE, R. L. (1967). The analysis and selection of test items. In D. Jackson & S. Messick (Eds.), *Problems in human assessment* (201–216). New York: McGraw-Hill.
WEDELL, D. H., PARDUCCI, A., & LANE, M. (1990). Reducing the dependence of clinical judgment on the immediate context: Effects of number of categories and type of anchors. *Journal of Personality and Social Psychology*, 58, 319–329.
WIGGINS, J. S., & GOLDBERG, L. R. (1965). Interrelationships Among MMPI Item Characteristics. *Educational and Psychological Measurement*, 25(2), 381–397.
WRIGHT, B. D., & MASTERS, G. N. (1982). *Ratings scale analysis*. Chicago, IL: MESA Press.
ZHU, W. (2002). A confirmatory study of Rasch-based optimal categorisation of a rating scale. *Journal of applied measurement*, 3(1), 1–15.
ZHU, W., & KANG, S. J. (1998). Cross-cultural stability of the optimal categorization of a self-efficacy scale: A Rasch analysis. *Measurement in Physical Education and Exercise Science*, 2(4), 225–241.
ZHU, W., TIMM, G., & AINSWORTH, B. A. (2001). Rasch calibration and optimal categorization of an instrument measuring women‘s exercise perseverance and barriers. *Research Quarterly for Exercise and Sport*, 72(2), 104–116.
ZHU, W., UPDYKE, W., & LEWANDOWSKI, C. (1997). Post-hoc Rasch analysis of optimal categorization of an ordered-response scale. *Journal of Outcome Measurement*, 1(4), 286–304.

VLIV KATEGORIZACE NA VNITŘNÍ STRUKTURU DOTAZNÍKU EXERCISE BARRIER INVENTORY

JAN ŠTOCHL, WEIMO ZHU

SOUHRN

Studie se zabývá vlivem množství kategorií odpovědí (3 a 5) na faktorovou strukturu dotazníku Exercise Barrier Inventory. Předchozí studie navrhly redukci počtu odpovědí z 5 na 3., což z pohledu statistiky znamená zmenšení informace a možný negativní dopad na validitu či reliabilitu dotazníku. Pomocí strukturálního modelování (resp. teorie položkových odpovědí) jsme nezjistili negativní vliv redukce na validitu tohoto dotazníku – faktorová struktura zůstala zachována.

Klíčová slova: Exercise Barrier Inventory, kategorizace odpovědí, validita, faktorová analýza, teorie položkových odpovědí

Jan Štochl, Ph.D.
stochl@ftvs.cuni.cz