Qual Quant (2009) 43:381–390
DOI 10.1007/s11135-007-9114-2

Fuzzy item response model: a new approach to generate membership function to score psychological measurement

Sen-Chi Yu · Berlin Wu

Received: 1 May 2006 / Accepted: 1 October 2006 / Published online: 12 June 2007

Abstract The aim of this study is to propose an new approach, fuzzy item response model (FIRM), which combines item response theory (IRT) and fuzzy set theory, in the educational or psychological measurement. Applying FIRM to improve the predictive validity of psychological measurement is verified. We set up a detailed procedure for the FIRM and apply it to a valuable empirical study with Beck Depression Inventory-II (Chinese version) administrated on outpatient diagnosed as depression was given. The results showed the correct classification of depression based on FIRM scoring was 80.3% while that of raw score was only 73.2%. That is, via FIRM scoring, 7.9% of the erroneous judgments of depression inferred from self-reported inventory were reduced. It is also suggested that considerable cost concerning prevention and cure of depression might be reduced via FIRM.

Keywords Fuzzy Item Response Model · Fuzzy set theory · Rasch model · depression

1 Introduction

In classical test theory (CTT), "Method of successive integral", or "raw score", is most straightforward and popular scoring method in psychological measurements (Guilford 1954; Yu 2005). In method of successive integral, alternatives listed in the scale is treated as equal-distance and scored as successive integral. For example, score of 4, 3, 2, or 1 was given if the alternative "strongly agree", "agree", "disagree", or "strongly disagree" was chosen, respectively. This scoring approach, however, has been criticized on the grounds that it is too

S.-C. Yu (✉)
Center for Teacher Education, Huafan University, No.1, Huafan Rd., Shiding Shiang, Taipei County 223, Taiwan ROC
e-mail: rhine@cc.hfu.edu.tw

B. Wu
Department of Mathematics, National Chengchi University, Taipei, Taiwan ROC
e-mail: berlin@math.nccu.edu.tw

simplistic (Nunnally and Bernstein 1994; Yu 2005). First, the assumption of equal-distance of adjacent alternative is questionable. Second, concerning the characteristic of variables, the descriptive terms applied in rating scales are linguistic variables rather than numerical variables (Yu 2005; Zimmermann 1996). Consequently, utilizing fuzzy set theory (FST) to score psychological measurement seems feasible. However, in contrast with the many engineering studies discussing fuzzy set theory, only a few such works have been published in psychological measurement.

Eliciting membership function is a critical concerning applying FST in scoring psychological measurement. In the fields of engineer, methods based on heuristic, probability to possibility transformations, histograms, nearest neighbor techniques, feed-forward neural networks, clustering, and mixture decomposition are utilized to generate membership functions (Medasani et al. 1998). Nonetheless, in the field of psychological measurement, few eliciting methods were proposed and applied. Some studies generating fuzzy numbers by CTT (classical test theory)-approach raw score (Wu 1995; Wu and Lin 2002a, 2002b; Lin 2001, 2003a, 2004; Law 1996) revealed that fuzzy set approach is more reliable and accurate than the raw scores. However, till now, the scaling and membership generating of these works using FST to psychological measurement are based on CTT rather than modern test theory, also called item response theory (IRT). In CTT approach, raw score is utilized to scale one's ability or trait. However, raw score suffices to accomplish a "meaningful measurement" (Wright 1999).

Psychological measurement must be unidimensional, linear, invariant and objective (Rasch 1960; Wright and Linacre 1989) to achieve a meaningful measurement (Wright 1997). Unidimensionality, the universal characteristic of all measurement, means that measurement of any object or entity describes only one attribute of the object measured (Thurstone 1931). Linearity refers to equality of measurement unit, that is, measurement must be linear allowing arithmetic operations to be performed on them. Invariance means that the measurement maintains its properties when the unit of measurement is changed. Objectivity means sample-free and test-free, that is, item calibrations must not depend on the respondent are used to estimate them, and measurements must not depend on which items are taken.

However, raw score is non-linear, and sample and test dependent. To attain objective, meaningful measurement, raw scores must be transformed into linear measures to enable subsequent analysis and inference. By contrast, IRT-based latent trait models meet the requirements of linearity, sample and test dependence and, consequently, predominate over CTT in psychological measurement.

The partial credit model (PCM, Masters 1982), a unidimensional IRT model for polychotomous items, parameterizes the difficulties of a series of categories, called "steps", in each item. A distinguished feature of PCM is that item steps may vary between items. The PCM is a member of the Rasch family of latent trait models with parameter separability, permitting "specifically objective" comparisons of person and items (Masters 1982; Masters and Wright 1984).

The previous paragraph suggests using IRT to generate fuzzy number seems to be feasible. Therefore, we propose a new scaling method, fuzzy item response model (FIRM), which uses PCM, an IRT-approach latent trait model, to construct fuzzy numbers and utilized these fuzzy numbers to score psychological measurements. Moreover, an empirical study is given to illustrate FIRM. Finally we compare the predictive validity of FIRM with that of raw score.

Table 1 Pass/fail scores for a three-step item i

Person j	Performance levels				Scores X_{ij}
	0 STRONGLY DISAGREE	1 DISAGREE	2 AGREE	3 STRONGLY AGREE	
1	1 → First Step	1 → Second Step	1	→ 1 Third Step	3
2	1 →	1 →	1		2
3	1 →	1			1
4	1				0

2 Methods

2.1 Introduction to partial credit model

Masters' (1982) partial credit model (PCM) is an application of Rasch's model for dichotomies. When an item provides only two scores 0 and 1, the probability of scoring 1 rather than 0 is expected increase with the ability being measured. In Rasch's model for dichotomies, the to probability of person j succeeding on item I is written as:

$$\frac{P_{ij1}}{P_{ij0} + P_{ij1}} = \frac{\exp(\theta_j - \delta_i)}{1 + \exp(\theta_j - \delta_i)} \tag{1}$$

where P_{ij1} is the probability of person j scoring 1 on item I, P_{ij0} is the probability of person j scoring 0, θ_j is the ability of person j, and δ_i is the difficulty of item I defined as the location on the measurement variable at which a score of 1 on item I is as likely as score of 0. The model is written here as a conditional probability to emphasize that it is a model for the probability of person j scoring 1 rather than 0.

In PCM, this expectation is modeled using the above-mentioned Rasch model for dichotomies above-mentioned:

$$\frac{P_{ijx}}{P_{ijx-1} + P_{ijx}} = \frac{\exp(\theta_j - \delta_{ix})}{1 + \exp(\theta_j - \delta_{ix})}, \quad x = 1, 2, \ldots, m_i \tag{2}$$

where P_{ijx} is the probability of person j scoring x on item I, P_{ijx-1} is the probability of person j scoring $(x - 1)$, θ_j is the ability of person j, and δ_{jx}, called a "step", is an item parameter governing the probability of scoring x rather than $x - 1$ on item i Table 1.

The interpretation of "step" is illustrated in Table 2.3. For an item on an attitude questionnaire, "completing the kth step" means choosing the kth response alternative over the $(k - 1)$th in response to the item. In Table 2.3, a person who chooses to AGREE can be considered to have chosen DISAGREE over STRONGLY DISAGREE (first step taken) and also AGREE over DISAGREE (second step taken), but to have failed to choose STRONGLY AGREE over AGREE (third step rejected).

The third step in item I listed in Table 2.2 is from level 2 to level 3. The difficulty of the third step governs how likely it is that a person who has already reached level 2 will complete the third step to level 3. Therefore, the probability of scoring 3 rather than 2 can be expressed as

$$\Phi_{ij3} = \frac{P_{ij3}}{P_{ij2} + P_{ij3}} = \frac{\exp(\theta_j - \delta_{i3})}{1 + \exp(\theta_j - \delta_{i3})}. \tag{3}$$

Likewise, the second step in item I is from level 1 to level 2 since a person cannot make a "0" by failing the second step. Therefore, the probability for a person making a "2" rather than "1" on item I is

$$\Phi_{ij2} = \frac{P_{ij2}}{P_{ij1} + P_{ij2}} = \frac{\exp(\theta_j - \delta_{i2})}{1 + \exp(\theta_j - \delta_{i2})}. \tag{4}$$

Similarly, the first step in item I is to make a "1" rather than a "0":

$$\Phi_{ij1} = \frac{P_{ij1}}{P_{ij0} + P_{ij1}} = \frac{\exp(\theta_j - \delta_{i1})}{1 + \exp(\theta_j - \delta_{i1})}. \tag{5}$$

Finally, as person j must make one of the four possible scores on item I,

$$\sum_{h=0}^{m_i} P_{ijh} = P_{ij0} + P_{ij1} + P_{ij2} + P_{ij3} = 1 \tag{6}$$

Equations 3–6 are readily solved to obtain one general expression for the probability of person j scoring x on item i:

$$P_{ijx} = \frac{\exp \sum_{k=0}^{x} (\theta_j - \delta_{ik})}{\sum_{h=0}^{m_j} \exp \sum_{k=0}^{h} (\theta_j - \delta_{ik})} \quad x = 0,\ 1, \ldots,\ m_i \tag{7}$$

2.2 Generating fuzzy numbers using FIRM

In this section, the procedures of generating fuzzy numbers using FIRM are demonstrated.

2.2.1 Traditional and fuzzy scoring

Fuzzy logic argues that membership degree that some individual belongs to a certain alternative (category) is a continuum value, gradual transition from 0 to 1, rather then a dichotomy, 0 or 1. According to this argument, in FIRM, subjects are free to choose more than one alternative for each item and, in turn, assign percentages on the chosen alternatives. The assigned percentages represent the degree of membership that some subjects belong to the category. Moreover, the sum of percentages of the chosen categories is restricted to 100%. Moreover, the triangular normal fuzzy numbers $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ were constructed to represent alternative 1 to 4, respectively.

Table 2 shows the examples of fuzzy scoring (FS) and traditional scoring. As shown in this table, the category assigned the most percentages is treated as the traditional scoring. If there are two most assigned categories, the former (the minimum of the two categories) will be taken as traditional scoring. The sum of fuzzy numbers multiplied by its membership degree, constitute the fuzzy scoring. Since the calculations of PCM require crisp number, the results of traditional scoring were utilized as crisp data for PCM algorithms. Whereas the results of fuzzy scoring, still fuzzy numbers, will be utilized for sequent analysis.

2.2.2 Fuzzy item response model

The procedures for generating fuzzy numbers using FIRM to scoring psychological measurements were as follows:

Step 1: Subjects are asked to choose and assign percentages on alternatives of items. The sum of assigned percentages, representing the membership degrees, in each item must be constrained to 100%.

Table 2 Examples of fuzzy and traditional scoring

	Assigned Percentages (Degree of Membership)%	Traditional Scoring (crisp value)	Fuzzy Scoring(FS) $FS = \sum \mu_{ijk}(\tilde{K})$ (interval value)
(a) Only One Alternative Chosen			
Alternative 1* ($\tilde{A}$)	100	**1**	$1 \times \tilde{A}$
Alternative 2 ($\tilde{B}$)	0		
Alternative 3 ($\tilde{C}$)	0		
Alternative 4 ($\tilde{D}$)	0		
(b) Two Alternatives Chosen			
Alternative 1* ($\tilde{A}$)	80	**1**	$0.8 \times \tilde{A} + 0.2 \times \tilde{B}$
Alternative 2 ($\tilde{B}$)	20		
Alternative 3 ($\tilde{C}$)	0		
Alternative 4 ($\tilde{D}$)	0		
(c) Two Most Assigned Alternatives			
Alternative 1* ($\tilde{A}$)	50	**1**	$0.5 \times \tilde{A} + 0.5 \times \tilde{B}$
Alternative 2 ($\tilde{B}$)	50		
Alternative 3 ($\tilde{C}$)	0		
Alternative 4 ($\tilde{D}$)	0		

* indicates the category assigned the most percentages

Step 2: Calculate the traditional scoring according to the procedures mentioned above.
Step 3: Calculate "step parameters" (δ_{ij}) defined in PCM as shown in Fig. 1. The PCM algorithm (Masters and Wright 1997) is shown in Eq. 7.
Step 4: Fuzzify crisp data into fuzzy data by constructing triangle fuzzy numbers using step parameters estimated in Step 3.

We try to map linguistic variables, Alternatives 1 to 4, into corresponding reasonable normal fuzzy numbers $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$, with triangular membership functions $\mu_{\tilde{A}}$, $\mu_{\tilde{B}}$, $\mu_{\tilde{C}}$ and $\mu_{\tilde{D}}$. These membership functions are shown in Fig. 1.

The x-axis represents ability, usually ranging from -3 to 3; while y-axis represents degree of membership, ranging from 0 to 1.

In Fig. 1, we first find the "step parameters" (δ_{ij}) estimated by PCM. We propose that subject with ability located between -3 and "step parameter 1" (δ_{i1}) will choose Alternative 1. For this reason, the triangular fuzzy number $\tilde{A} = (-3, (-3+\delta_{i1})/2, \delta_{i1})$ with -3 and δ_{i1} being the lower and upper bounds, respectively, and $(-3+\delta_{i1})/2$ as the most likely value for $\tilde{A}$. In Fig. 1, we draw a line segment from $((-3+\delta_{i1})/2, 1)$ to $(-3, 0)$ and $(\delta_{i1}, 0)$ to characterize the membership of function of $\tilde{A}$.

Next, we propose that subject with ability located between "step parameter 1" (δ_{i1}) and "step parameter 2" (δ_{i2}) will choose Alternative 2 and the middle point between these two step parameters should receive the maximum degree of membership. Therefore, the triangular fuzzy number $\tilde{B} = (\delta_{i1},\ (\delta_{i1}+\delta_{i2})/2,\ \delta_{i2})$ with δ_{i1} and δ_{i2} being the lower and upper bounds, respectively, and $(\delta_{i1}+\delta_{i2})/2$ being the middle point which is the most likely value for $\tilde{B}$. In Fig. 2, we draw a line segment from $((\delta_{i1}+\delta_{i2})/2, 1)$ to $(\delta_{i1}, 0)$ to represent the left leg and another line segment from $((\delta_{i1}+\delta_{i2})/2, 1)$ to $(\delta_{i2}, 0)$ to represent the right leg of the triangular fuzzy number.

Likewise, we proposed $\tilde{C} = ((\delta_{i2},\ (\delta_{i2}+\delta_{i3})/2,\ \delta_{i3})$ and $\tilde{D} = (\delta_{i3},\ (\delta_{i3}+3)/2, 3)$ to characterize Alternatives 3 and 4, respectively.

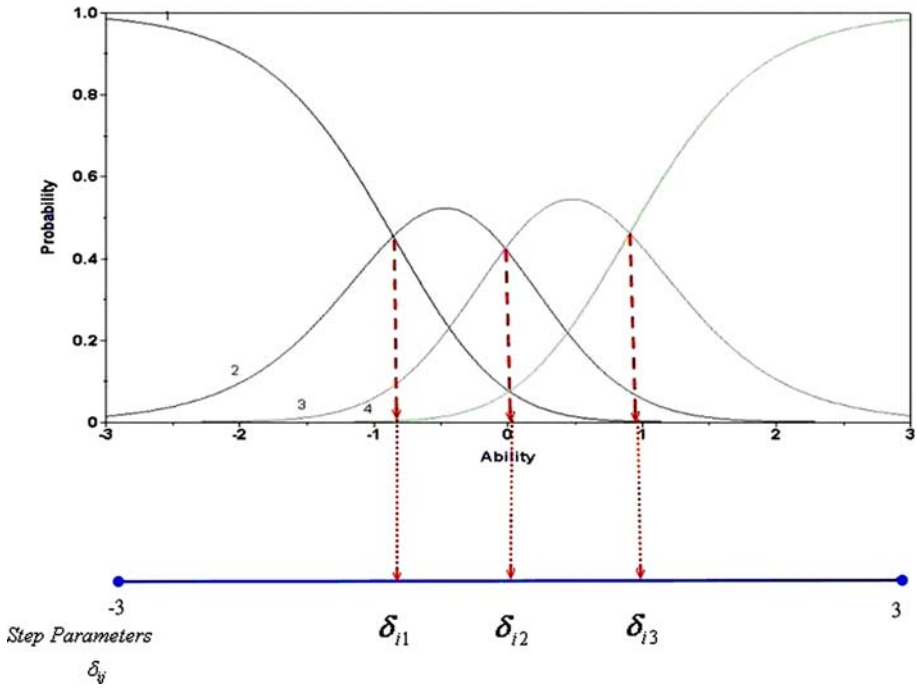

Fig. 1 Calculate "step parameters" (δ_{ij}) via PCM

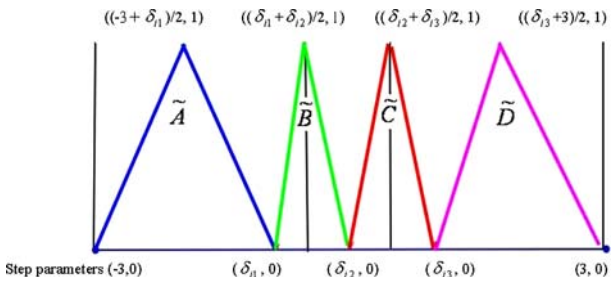

Fig. 2 Generating triangular fuzzy numbers via step parameters

2.2.3 Scoring of FIRM

Till Step 4 of FIRM, fuzzy score of single item was denoted by a triangular fuzzy number rather than crisp number. To calculate the aggregate fuzzy score (AFS) to represent the score of complete psychological measurement, the operations of fuzzy number were utilized in FIRM.

The algebraic operations for triangular fuzzy number were listed as follows (Chen and Hwang 1992):

Let $\tilde{M}$, $\tilde{N}$ be two triangular fuzzy numbers:

$$\tilde{M} = (m, \alpha, \beta), \quad \tilde{N} = (n, \gamma, \delta)$$

Then

$$\text{Addition}: \quad \tilde{M}(+)\tilde{N} = (m+n,\ \alpha+\gamma,\ \beta+\delta)$$
$$\text{Subtraction}: \quad \tilde{M}(-)\tilde{N} = (m-n,\ \alpha+\delta,\ \beta+\gamma)$$

For instance, subject j completed a three-item scale. The scoring of the three items were denoted as triangular fuzzy numbers $i_1 = (0, 1, 2)$ $i_2 = (1, 2, 3)$, and $i_3 = (0, 1, 2)$ respectively. Consequently, the aggregate fuzzy score (AFS), still a fuzzy number, was:

$$\text{AFS} = (0, 1, 2) + (1, 2, 3) + (0, 1, 2) = (1, 4, 7)$$

From sequent statistical calculation, AFS was defuzzified into crisp number using the center of gravity (COG) method. COG calculates the center of gravity of the support of the fuzzy number weighted by the membership grade. The center of gravity of fuzzy set $\tilde{X}$ with membership function $\mu_{\tilde{A}}$, $\text{GR(X)} = \int_{-\infty}^{\infty} x\mu_x(x)\mathrm{d}x / \int_{-\infty}^{\infty} \mu_x(x)\mathrm{d}x$

For a triangular fuzzy number $\tilde{X}(\text{a, b, c})$, $\text{GR(X)} = (\text{a} + \text{b} + \text{c})/3$, (Zimmermann 1996). The defuzzified AFS, called total fuzzy score (TFS) were used for sequent statistical analysis.

2.3 Empirical example: instrument and sample

The total sample used in this study consisted of participants recruited from two separate populations: (a) outpatients of a psychiatric clinic who were diagnosed as suffering from depression as the depressed sample, and (b) undergraduates as the non-depressed sample.

A total of 240 subjects were selected from outpatients who visit the psychiatric clinic at Taipei Municipal Hoping Hospital during July and August in 2004 and were diagnosed as having depression symptoms. The self-reported instrument utilized in this study was administrated by the researcher while the severity of depression was diagnosed by a psychiatrist. Since depression symptoms may appear in many mental disorders, the outpatients who were diagnosed with depression in this study include the following disorders: Major Depression Disorder, Bipolar Disorders, Dysthymic Disorder, and Adjustment Disorder with Depressed Mood.

A total of 321 students from several educational psychology and educational test and evaluation classes in Taiwan were recruited in this study as non-depressed sample.

The instrument in this study was the Chinese version of Beck Depression Scale II (C-BDI-II).BDI-II (BDI, Beck et al. 1996), the most widely used and investigated self-report measure of depression for clinic samples, is a self-reported instrument for measuring the severity of depression in adolescents and adults through items showing varying degrees of the main cognitive, affective, and physiological aspects of clinical depression. The C-BDI-II was adapted from the original BDI-II by the Chinese Behavioral Sciences Society and made available in 2000.

A number of studies have generally found that the BDI–II has high internal consistency (alpha coefficient $> .90$) and moderate to strong convergent validities with other self-reported measures, such as CES-D, the Reynolds Adolescent Depression Scale, and clinical rating scales of depression in adult and adolescent psychiatric patients, college students, and normal adults (Krefetz et al. 2002).

Table 3 Classification Table

		Predicted Group Non-depressed	Depressed	Percentage Correct
(a) FIRM				
Observed Group	Non-depressed	272	46	85.5
	Depressed	64	175	73.2
Overall Percentage				80.3
(b) Raw Score				
Observed Group	Non-depressed	261	58	81.8
	Depressed	96	143	59.8
Overall Percentage				72.4

3 Results

Predictive validity was employed in this study to investigate the validity of FIRM. Predictive validity refers to the relationship between a predictor and a criterion event (Nunnally and Bernstein 1994). In this study, two different scoring schema, raw scores and FIRM, yielded two different predictors whereas suffering from depression as diagnosed by psychiatrist served as the criterion. Logistic regression was applied to investigating the relation between scoring schemas and diagnosis of suffering from depression (binary outcome).

Since explanatory variables in this study were continuous, the Hosmer-Lemeshow statistic, instead of chi-square statistic, was employed to measure goodness-of-fit (Agresti 1996).

Concerning FIRM, the Hosmer-Lemeshow statistics equaled 12.121, with df = 8, p-value = 0.146, indicating a good fit. The analytical results showed that the probability of correct classification of depressed and non-depressed was 80.3%.

Concerning raw score, the Hosmer-Lemeshow statistics equaled 29.368, with df = 8, p-value = 0.1146, indicating bad fit. The analytical results showed that the probability of correct classification of depressed and non-depressed was only 72.4%. Obviously, the predictive validity of raw score is inferior to that of FIRM. These findings reveal that FIRM, compared with raw scores, yields better model fit and more accurate estimation for predicting depression.

Since depression is one of the most threatening factors concerning human health in the 21st century, the screening, diagnostic evaluation and treatment of depression are important issue of public health. Psychological inventories are most straightforward and economical tools for screening depression, therefore, a valid scoring schema is essential to achieve an accurate prediction. As shown in this study, the prediction based on FIRM is more accuracy than that of raw score, since 7.9% miss-classification of depression was reduced Table 3.

4 Discussion

The traditional scoring method of psychological measurement, raw score, had been criticized by advocates of IRT for non-linearity and sample- and item-dependence. (Nunnally and Bernstein 1994; Yu 2005). Besides, the alternatives listed in psychological measurement are linguistic variables in nature; therefore, this study proposed a scoring method, FIRM, integrates PCM, a one-parameter IRT model, with FST to scoring psychological measurements. The empirical example concerning Chinese Version of Beck Depression Scale II administrated on outpatients with depression and non-depressive college students showed

that FIRM, compared with raw score, exhibited more accurate prediction on depression. The probability of correct classification of depression and non-depression was improved 7.9%, definitely. That is, via FIRM, 7.9% of the erroneous judgments of depression inferred from self-reported inventory were reduced. Concerning depression costs the US $43.7 billion a year in medical expenses and lost productivity (Goleman 1993), not to mention the cost in human suffering cannot be estimated. This study showed that, FIRM which applies IRT to generating membership to scoring psychological measurement, is a more accurate scoring schema than raw score. Via FIRM, erroneous judgments of depression were reduced and medical cost concerning depression was reduced.

The analytical results also support that fuzzy logic conveys human thinking more accurately than can crisp logic. Theoretically, fuzzy logic should handle vagueness and imprecision in human thinking better traditional crisp logic. Empirically, the validity investigation in this study proves the above theory. When vagueness in human thinking is inevitable, crisp (binary) logic, which describes real-world situations using a simplified mathematical model, trades accuracy for simplicity (Kosko 1993). Thus, when selecting one alternative among many statements presented in rating scales to describe a person's mood state or attitude, some force fitting and rounding off are inevitable, resulting in some loss of information. Therefore, the FIRM scheme, based on fuzzy logic, is a valid scoring schema for psychological measurement.

Acknowledgements The authors would like to thank National Science Council of Republic of China for the support given to this study. (Project Number: NSC95-2413-H-211-001)

References

1. Agresti, A.: An introduction to categorical data analysis. John Wiley and Sons, New York (1996)
2. American Psychiatric Association.: Diagnostic and statistical manual of mental disorders. American Psychiatric Association, Washington DC (1994)
3. Beck, A.T., Steer, R.A., Brown, G.K.: Manual for the beck depression inventory–II. Psychological Corporation, San Antonio, TX (1996)
4. Bilgic, T., Turksen, B.: Measurement of membership function: theoretical and empirical work. In: Dubios, D., Prade, H. (eds.) Fundamentals of fuzzy sets, pp. 195–232.Kluwer, Boston (2000)
5. Cheng, S.J., Hwang, C.L: Fuzzy multiple attribute decision making: methods and applications. Springer-Verlag, Berlin (1992)
6. Dubios, D., Ostasiewicz, W., Prade, H.: Fuzzy sets: history and basic notions. In: Dubios, D., Prade, H. (eds.) Fundamentals of fuzzy sets, pp. 21–124.Kluwer, Boston (2000)
7. Goleman, D. Costs of depression are on a par with heart disease, a study says. The New York Times, December 3, 1993
8. Guilford, J.P.: Psychometric methods. 2nd ed. McGraw-Hall, New York (1954)
9. Hopkins, K.D., Stanley, J.C., Hopkins, B.R.: Education and psychological measurement and evaluation. Prentice Hall, Englewood Cliffs, NJ (1990)
10. Krefetz, D.G., Steer, R.A., Gulab, N.A., Beck, A.T.: Convergent validity of the Beck Depression Inventory–II with the Reynolds adolescent depression scale in psychiatric inpatients. J. Person. Assess. **78**(3), 451–460 (2002)
11. Klir, G., Yuan, B.: Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, Englewood Cliffs, NJ (1995)
12. Kosko, B.: Fuzzy thinking: the new science of fuzzy logic. Hyperion, New York (1993)
13. Law, C.K.: Using fuzzy numbers in educational grading system. Fuzzy Sets Syst. **83**, 311–323 (1996)
14. Lin, Y.H.: The simulation study of reliability of fuzzy linguistic scales. J. Test Stat. **9**, 193–219 (2001)
15. Lin, Y.H.: The comparison of different algorithms and scoring. J Taichuang Teacher Coll. **17**(2), 279–304 (2003)
16. Lin, Y.H.: The data simulation and empirical study of fuzzy linguistic scoring. (NSC92-2413-H-142-004) (2004)

17. Linacre, J.M.: A User's guide to WINSTEPS Rasch-model computer programs. Winsteps, Chicago (2005)
18. Masters, G.N.: A Rasch model for partial credit scoring. Psychometrika **47**, 149–174 (1982)
19. Masters, G.N., Wright, B.D.: The essential process in a family of measurement models. Psychometrika **49**, 529–544 (1984)
20. Masters, G.N., Wright, B.D. : The partial credit model. In: van der Linden, W.J., Hambleton, R.K. (eds.) Handbook of modern item response theory, pp. 101–122.Springer-Verlag, New York (1997)
21. Medasani, S., Kim, J., Krishnapuram, R.: An overview of membership function generation techniques for pattern recognition. Int J Approx Reason **19**, 391–471 (1998)
22. Nguyen, H., Wu, B.L.: Fuzzy mathematics and statistical analysis. Junjie, Taipei (2000)
23. Nunnally, J.C., Berstein, I.H.: Psychometric theory. 3rd ed. McGraw-Hall, New York (1994)
24. Rasch, G.: Probabilistic models for some intelligence and attainment tests. [Danish Institute of Educational Research 1960, University of Chicago Press 1980, MESA Press 1993] MESA Press, Chicago (1960)
25. Thurstone, L.L.: Measurement of social attitudes. J. Abnormal Soc. Psychol. **26**, 249–269 (1931)
26. Wright, B.D.: Rasch factor analysis. Rasch Measure Trans **8**(1), 348–349 (1994)
27. Wright, B.D.: Fundamental measurement for psychology. In: Embretson, S.E., Hershberger, S.L. (eds.) The new rules of measurement: what every educator and psychologist should know, pp. 235–255. Lawrence Erlbaum, Hillsdale, NJ (1999)
28. Wright, B.D., Linacre, J.M.: Observations are always ordinal: measures, however, must be interval. Arch. Phys. Med. Rehabilitation **70**, 857–860 (1989)
29. Wright, B.D., Masters, G.N.: Rating scale analysis: Rasch measurement. MESA Press, Chicago (1982)
30. Wu, B.L.: Fuzzy statistics analyzing: a new approach to survey research. Newsletter Nat. Chengchi Univ. **2**, 65–80 (1995)
31. Wu, B.L., Hsu, Y.Y.: The use of kernel set and sample memberships in the identification of nonlinear time series. Soft Comput **8**, 207–216 (2004)
32. Wu, B.L., Lin, Y.H.: Fuzzy mode and its applications in educational and psychological assessment analysis. Paper Presented on the Second International Conference on Information. China, Beijing (2002a)
33. Wu, B.L., Lin, Y.H.: The introduction of fuzzy mode and its applications. J. Test Stat. **47**, 23–27 (2002b)
34. Yu, S.: Fuzzy partial credit scaling: applying fuzzy set theory to scoring rating scales. Unpublished doctoral dissertation, National Chengchi University, Taipei, Taiwan (2005)
35. Zadeh, L.A.: Fuzzy sets. Inform Control **8**, 338–352 (1965)
36. Zimmermann, H.J.: Fuzzy set theory and its applications. Kluwer, Boston (1996)