Asia Pacific Journal of Educators and Education, Vol. 27, 119–138, 2012

VALIDATING STUDENT-PARENT ACTIONS MODEL (SPA) OF CONSISTENCY USING THE RASCH MEASUREMENT MODEL

William Koh Siaw Yen[1]* and Ong Saw Lan[2]
[1]SMK Bandar Tasek Mutiara
14120 Pulau Pinang
[2]School of Educational Studies
Universiti Sains Malaysia
11800 USM Pulau Pinang
*Corresponding author: william8m@yahoo.com.my

Abstract: The objective of this paper is to investigate the consistency of item selection and the existence of unidimensionality within each construct of the Student-Parent Actions Model (SPA) measurement scale using item response theory (Rasch measurement model). Based on Rasch model, which uses the WINSTEPS program, 373 students' responses were analysed. The sample group, which had a mean age of 16.41 years, was selected from five secondary schools through convenience sampling. The initial 62 items of SPA were analysed within the eight 'a priori' hypothesised dimensions. Rasch's misfit statistics were used to eliminate items that violated the infit and outfit MNSQ criteria of being between 0.70 and 1.30. In combination with the first principal component analysis of standardised residual (PCAR), Rasch's fit statistics were used to analyse unidimensionality within each individual construct. From the Rasch analysis results, 35 items were selected. In total, 27 items were similar to the SPA selected through structural equation modelling (SEM) analysis. All eight constructs exhibited unidimensionality, as PCAR analysis showed explained variances ranging from 47.10% to 73.6% for each individual construct. However, the eigenvalues of unexplained variances in the first contrast were all less than 3.0 indicating unidimensionality within each construct. Overall, SPA exhibited high consistency in the selection of items between the Rasch and SEM methods. When compared to the original SPA, the high percentage of similarity (approximately 80%) for the selected items indicated item selection stability for SPA across the two methods. Multidimensionality for the instrument was also detected.

Keywords: Rasch Measurement Model, structural equation modelling, item selection

Abstrak: Objektif kajian adalah untuk menyelidik konsistensi dalam pemilihan item dan kewujudan unidimensionaliti untuk konstruk-konstruk dalam *Student-Parent Actions Model* (SPA) menggunakan teori item respons (Model Pengukuran Rasch). Respons daripada 373 pelajar dianalisis berdasarkan model Rasch dengan menggunakan program WINSTEPS. Min umur sampel adalah 16.41 tahun dan dipilih menggunakan kaedah *convenience sampling* daripada lima buah sekolah menengah. Enam puluh dua item yang merangkumi lapan konstruk yang dihipotesiskan *a priori* dianalisa. Statistik Rasch digunakan untuk pemilihan item. Kriteria yang digunakan untuk memilih item ialah MNSQ 0.70–1.30. Statistik Rasch yang digabungkan dengan PCAR digunakan untuk menganalisis unidemensionaliti dalam setiap konstruk. Tiga puluh lima item dipilih berdasarkan kriteria Rasch sementara 27 item dipilih apabila menggunakan kaedah *structural equation modelling* (SEM). Kesemua konstruk menunjukkan unidimensionaliti apabila nilai PCAR berada dalam lingkungan 47.10% hingga 73.60%. Nilai eigen kurang daripada 3.0 turut memberikan indikasi bahawa konstruk adalah unidimensi. Secara keseluruhannya, SPA mempamerkan konsistensi yang tinggi walaupun menggunakan dua kaedah yang berbeza di antara Rasch dan SEM. Persamaan yang mencapai 80% di antara item terpilih memberikan implikasi bahawa pemilihan item adalah agak stabil walaupun menggunakan dua kaedah berbeza. Akhir sekali, multidimensionaliti instrumen turut dikesan.

Katakunci: Model Pengukuran Rasch, Permodelan Persamaan Struktural, pemilihan item.

INTRODUCTION

Student-Parent Actions Model (SPA) was developed using classical test theory (CTT) and structural equation modelling (SEM) to investigate the contributing dimensions of parents' observable actions towards their children's academic achievement and discipline in Malaysia (Koh & Ong, 2010). SPA is a self-report questionnaire that captures parental action constructs related to their adolescent children's academic achievement. Rasch measurement model (Rasch) and SEM are two approaches increasingly employed in all fields of social science research (Bartholomew & Knott, 1999; Bollen, 2002; Lamoureux, Pallant, Pesudovs, Hassel, & Keeffe, 2006; MacCallum & Austin, 2000; Merrell & Tymms, 2005; Pallant & Tennant, 2007). Some authors have suggested that using the Rasch measurement model to assess scaling properties of questionnaires is preferable to using CTT and confirmatory factor analysis (Nijsten, Unaeze, & Stern, 2006; Prieto, Alonso, & Lamarca, 2003; Wright, 1999). Others have claimed that these two approaches exhibit similar statistical frameworks and support hypothesis-driven data analysis in the social sciences (Bartholomew & Knott, 1999; Glockner-Rist & Hoijtink, 2003) and have therefore suggested integrating both approaches. However, Waugh and Chapman (2005) commented that factor analysis may provide misleading evidence that a scale is working well when it is not. Therefore, this study is interested in reassessing the item selection consistency (Allen & Oshagan, 1995; Bollen, 1989) of SPA when Rasch model approach is utilised.

Hence, the objectives of this study are (1) to determine the consistency of item selection in SPA and (2) to use Rasch analysis to validate the unidimensionality of the model's individual constructs.

The original SPA contains 34 items and uses a 5-point Likert-type scale. Items are grouped into eight dimensions, which are hypothesised 'a priori' using the SEM technique. Example items include the following: (1) "My parents discuss my future career with me", (2) "My parents check my homework after it has been done", (3) "My parents restrict my viewing time when I watch too much television", and (4) "When I do not get very good grades, my parents will encourage me to try harder". For the 34-item scale's score, the Cronbach Alpha was 0.96. Based on SEM, SPA exhibits evidence of reliability and construct validity. Convergent validity was observed based on the high factor loadings of the items and acceptable goodness of fit indices (Koh & Ong, 2010). Discriminant validity was evidenced in the low correlation among the SPA constructs.

The Rasch technique was chosen because it provided information regarding the extent to which each item was difficult to endorse (i.e., whether some SPA items are more difficult to endorse than others). It was also helpful to know whether the items' difficulty levels reflected the full range of the respondents' trait levels.

Developing instruments using different statistical tools may produce different results. These differences are comprehensible, as the techniques and underpinning theories are basically different. In addition, Rasch analysis, which is a sophisticated approach to questionnaire development that utilises modern psychometric methods (Lamoureux et al., 2006), converts categorical data into a continuous scale. Rasch converts ordinal scale into interval-based measures using logarithm functions (logit). In addition to providing estimates of item and person measures on an interval scale, Rasch also calculates item difficulty in relation to students' achievement. However, this is not available in SEM. SEM assumes that the ordinal Likert-scale categories are divided into equal parts that align with the strengths of the categories. Therefore, the responses are assumed to be scaled as intervals. Each method inevitably has its strengths and weaknesses.

LITERATURE REVIEW

SPA was developed using a structural equation modelling approach (see Koh & Ong, 2010). The items were created based on a review of existing theories and research. The main theory integrated into this instrument was ecological system theory (Bronfenbrenner, 1979), which puts forward the idea that entities closest to a child will have greater influences on the child. However, various measurement development theories may produce different end results for the instrument created. Previous results have suggested combining the two approaches (SEM and Rasch) to complement each other (Glockner-Rist & Hoijtink, 2003; Griffin, 2005). In this study, the final outcome of items selected based on Rasch was compared to the original SPA. The theories upon which SEM and Rasch analysis are based are briefly discussed.

Rasch can be generalised to polytomous items with ordered categories. This includes the partial credit model (PCM) (Masters, 1982) and the rating scale model (RSM) (Andrich, 1978), which can be expressed mathematically as:

$$P_{xni} = \frac{\exp(\sum_{j=1}^{x}(B_n - \delta_{ij})}{\sum_{k=1}^{ni}(\exp\sum_{j=1}^{k}(B_n - \delta_{ij})}$$

The variable (δ_{ij}) denotes the item step difficulty, B_n is the trait parameter of person n and P_{xni} represents the probability of an endorsed response. In the model, the calibration processes involve comparing the amount of the trait (B_n) possessed by a person and the amount of the trait demanded by the item (δ_{ij}). Generally, the higher the value of the item step difficulty (δ_{ij}), the greater the difficulty level of a particular step relative to other steps within an item (Embretson & Reisse, 2000).

RSM is actually a subset of PCM. In RSM, the step structure is restricted to the same structure for all items; thus, a common set of δ is estimated. This parameter δ is also known as the threshold (Andrich, 1978). Thus, RSM is only useful when distances between categories are equal for all items. In Rasch, item level difficulty is placed on a common metric parallel to the person's latent trait. This allows for comparisons between items and persons. The analysis was applied to determine whether the rating scale was used in the expected manner (e.g., students with high academic achievement would be expected to use higher item ratings, whereas students with lower academic achievement would be expected to use lower item ratings). The Rasch measurement model provides a connection between a person's total score and the items of the instrument by placing the student's measure and item measure on the same linear continuum. Rasch is based on minimising the residuals between the predicted and observed location parameters of items (and persons) on a latent variable. Furthermore, by using Rasch analysis, which converts ordinal scale into interval-based measures (log-odd metric or logit), the issue of scale metrics can be resolved. Item goodness of fit statistics may be used to determine the extent to which each item fits the construct it will be used to measure. Hence, the fit statistics permit assessment of the validity of the overall measure by providing a means to identify poorly functioning and/or biased items (Green & Frantom, 2002). Item fit can be interpreted as an index that reflects the convergence of the items to the construct being measured.

Although there is no definitive rule for fit cut-off values, some suggestions can be adopted, including the following:

1. Mean square (infit and outfit) between 0.6 and 1.4 (Bond & Fox, 2001)
2. Mean square (infit and outfit) between 0.8 and 1.2 (Bode & Wright, 1999)
3. Mean square (infit and outfit) of less than 1.3 for samples < 500, 1.2 for 500–1000 and 1.1 for samples > 1000 (Smith, Schumacker & Bush, 1998)

The mean standardised infit and outfit for the person and item are expected to be 0.0. If the value is negative, it indicates overfit. In other words, the data fit the model better than the expected fit, and this may signal some redundancy, which may be due to irrelevant items. Next, the separation index for items (G_i) and the separation index for person (G_p) were examined. Person separation of $G_p > 1.0$ indicates that the measurement is on a continuum and that there is sufficient breadth in position. Item separation of $G_i > 1.0$ indicates a broader continuum of rating spread. Normally, G_i will be larger than G_p because the number of items is normally smaller compared to the number of samples. The related formulae are shown below:

$$G_p = \frac{SA_p}{SE_p}$$

$$G_i = \frac{SA_i}{SE_i}$$

SA denotes the standard deviation, and *SE* is the average measurement error. The subscript *p* stands for person, and the subscript *i* represents an item. Another alternative index is the person separation reliability, which is a conceptual internal consistency of persons in rating the items. In combination with principal component analysis of standardised residual (PCAR), the unidimensionality of the construct could be identified (Karabatsos, 2000). In this analysis, each dimension was individually examined and confirmed for its unidimensionality properties. According to Bond and Fox (2001), the fit indices of each item provide indicators of how well each item fits within the underlying construct. Hence, these can ascertain whether the assumption of unidimensionality holds up empirically as well. The Rasch model is used to identify a series of items and scales with stable measurement properties. Item misfit indicates that an item must be excluded from the scale.

However, SEM is based on minimising the residuals between predicted and observed correlations in maximising a hypothesised relationship between the items and a latent variable. SEM is based on analysing variances and covariances among indicators (observed variables) and latent variables (unobserved variables). Various parameters are estimated, and goodness-of-fit indices are considered in making decisions about the instrument's accuracy. In other words, SEM is able to determine how well the data fit the theoretical model.

Research Questions

Because different quantitative techniques may produce different sets of instrument content due to their respective underlying theories, it is important to look at the consistency of item selection for SPA. Therefore, the research questions posed in this study are as follows:

1. To what extent are the items selected for SPA consistent across the Rasch and structural equation modelling methods?
2. Do the individual constructs in SPA exhibit unidimensionality?

METHODOLOGY

The processes undertaken in this study are simplified graphically in Figure 1. The data collected in the initial study for the original SPA were reused and analysed using Rasch measurement model. In total, 373 secondary-school students (160 males and 213 females) with a mean age of 16.41 years were involved in the study. They were selected from five secondary schools through convenience sampling. Items were eliminated based on a few criteria provided by the WINSTEPS software. One set of such criteria is the misfit of items, which is drawn from infit and outfit parameters.

In combination with PCAR, the unidimensionality of the construct can be identified (Karabatsos, 2000). Infit values between 0.70 and 1.30 were applied, as they were generally regarded as acceptable for samples less than 500 (Smith, Schumacker, & Bush, 1999). If the person's mean measure was less than –1.0, the items were considered as potentially too difficult. However, if it was more than 1.0, the items were considered too easy for the sample, and a revision of the questionnaire was possibly necessary. The mean standardised infit and outfit (ZSTD) were expected to be 0.0. The standard deviation (SD) of the standardised (ZSTD) infit is an index of overall misfit for persons and items (Bode & Wright, 1999). The standard deviation of the standardised infit (ZSTD SD) has a cut-off maximum value of |2.00| (Bode & Wright, 1999; Wright & Masters, 1982). The separation index (SEPAR) measures the spread of both items and persons in standard error units. It can also be thought of as the number of levels into which the sample of items and persons can be separated.

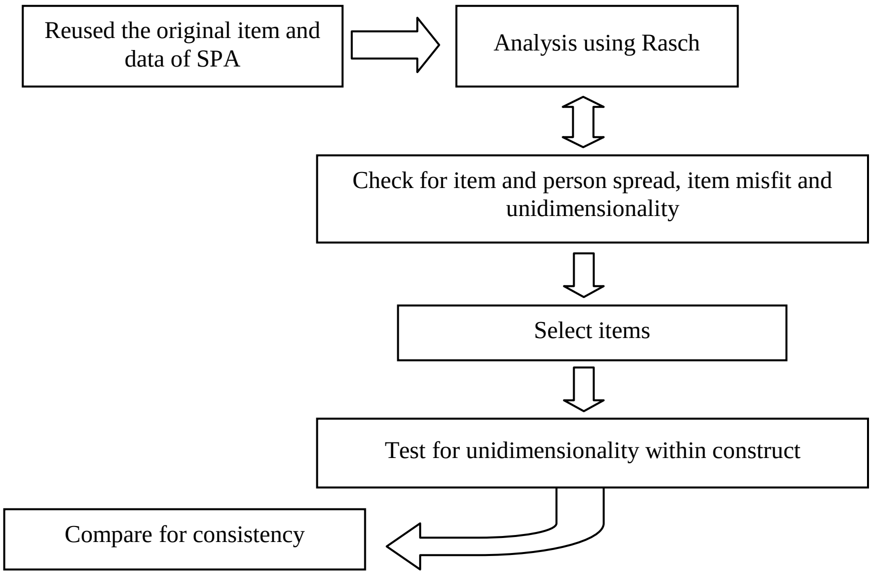

Figure 1. The phases involved in the SPA study

To fulfil the condition of unidimensionality, unexplained variance in PCAR must have a strength of an eigenvalue < 3 (Linacre, 2006). In this analysis, each dimension was individually examined for its unidimensionality property after the groups of items representing each construct were selected. According to Bond and Fox (2001), the fit indices of each item also provide indicators of how well each item fits within the underlying construct. Hence, these may ascertain whether the assumption of unidimensionality also holds up empirically. The Rasch model is used to identify a series of items and scales with stable measurement properties. Item misfit is an indication that an item needs to be excluded from the scale.

DATA ANALYSIS

All 'a priori' dimensions were tested for their convergent properties within the underlying conceptual meaning. Rasch was used consecutively to analyse the items for all eight constructs of SPA.

The overall model fit results for the 62 items and 373 persons are shown in Table 1.

Table 1. Overall model fit information, separation and mean measure of SPA

	MEAN	INFIT		ZSTD SD	OUTFIT		ZSTD SD	SEPAR	REL
		MNSQ	ZSTD		NSQ	ZSTD			
PERSON	0.37	1.03	–0.1	2.1	1.05	0.0	2.0	4.47	0.95
ITEM	0.00	1.00	–0.2	3.2	1.05	0.1	3.5	8.31	0.99

Note. SEPAR = Separation, REL = reliability

Rasch analysis showed that the mean measure for person was higher than 0.00; it yielded a value of 0.37. For these data, the person infit and outfit MNSQ was 1.03 and 1.05, respectively, and the item infit and outfit MNSQ was 1.00 and 1.05, respectively. As shown in Table 1, the person infit and outfit ZSTD were –0.1 and 0.0, respectively. The item infit and outfit ZSTD were –0.2 and 0.1, respectively. By examining the standard deviation of the standardised infit (ZSTD SD), the indices showed violation of the cut-off value of |2.00| (Bode & Wright, 1999; Wright & Masters, 1982). The person infit ZSTD SD was 2.1, and the item infit ZSTD SD yielded a value of 3.2.

Unidimensionality of the latent trait is determined by examining the first contrast from the items' PCAR. In the unexplained variance of the first contrast, a "secondary dimension" exists if it has the strength of at least three items. The results of PCAR (Table 2) indicate that SPA is not unidimensional, which supports the 'a priori' judgement that SPA is multidimensional. The unexplained variance in the first contrast is 4.6.

Table 2. Contrast 1 from PCAR

	Eigenvalue	Empirical (%)	Modelled (%)
Total raw variance in observations	104.7	100.00	100.00
Raw variance explained by measures	42.7	40.80	9.70
Raw variance explained by persons	15	14.30	4.00
Raw variance explained by items	27.6	26.40	5.70
Raw unexplained variance (total)	62	59.20	60.30
Unexplained variance in first contrast	4.6	44.40	

Next, each individual construct was analysed. Items were retained based on the accepted infit/outfit MNSQ values ranging from 0.70 and 1.30, and the observed values of item data points, which are higher than their expected values. SPA includes eight dimensions: Aspiration (ASP), Homework (HWK), Conduciveness (COND), Religiosity (REL), Control (CONT), Motivation (MOV), Warmness (WARM) and Conflict (CONF).

Aspiration (ASP) Dimension

The first dimension examined was Aspiration (ASP). ASP is defined as parents' hopes for their children's futures (Koh & Ong, 2010). The item fit statistics were used to present overall model fit information (see Table 3). The analysis showed that overall ASP item infit MNSQ = 1.00 and outfit MNSQ = 0.99, which were within the stipulated criteria of being between 0.70 and 1.30. However, the means of the standardised infit and outfit, which are best at 0.00, were –0.1 and –0.2, respectively. On average, these values are an indication of good fit. The standard deviation of the standardised infit is an index of overall misfit for persons and items (Bode & Wright, 1999). The standardised infit standard deviation was violated with a value of 3.2 (accepted value is < 2.0), indicating item misfit. The item separation for this case was 9.68, which implied a broad continuum of measurement. Finally, item reliability was 0.99. The misfit order was then inspected (see Table 4). The infit and outfit MNSQ were within the set criteria. However, Items A2, A7 and A6 did exhibit higher values of infit and outfit MNSQ. The percentage of data points within 0.5 score points of their expected values were 55.6% (Item A2), 45.5% (Item A7) and 61.5% (Item A6), which were below the expected match prediction. Furthermore, the three items exhibited the lowest point biserial measures and ranged from 0.62 to 0.70. These results supported the decision to delete the three items.

To test for unidimensionality within the ASP construct, PCAR was conducted (see Table 5). It showed that the items fulfilled the unidimensionality criteria, as there was no sign of a secondary dimension (the eigenvalue of the unexplained variance in the first contrast < 3). The four items explained a total of 65.10% of the variance, which is considered high.

Homework (HWK) Dimension

HWK refers to parents' involvement with their children's homework. The item standardised infit and outfit measures for HWK were –0.5 and –0.4, respectively. The standardised infit standard deviation, which must be < |2.0|, was violated with a value of 2.9, indicating item misfit. The item separation for this case was 10.24, which implied a broad continuum of measurement. The item reliability was 0.99. The infit standardised SD was 2.9, and it exceeded the cut-off value of |2.0|. Hence, Rasch confirmed the need for item deletion. Following the item misfit order (see Table 4), Items H3, H1, H6, and H7 were arranged in descending order for deletion. Item H3 exhibited infit and outfit values exceeding the criterion of 1.30. An infit value of 1.62, which was above 1.30, indicated a lack of unidimensionality for the HWK dimension; this conclusion was also supported with moderate point biserial correlation (0.48). This finding may indicate random response or may indicate that the item was either endorsed or not endorsed by the respondents. However, the infit and outfit MNSQ of H1, H6 and H7 were within the set criteria. The percentage of data points within 0.5 score points of their expected values was 40.7% (Item H3), which was below the expected match prediction of 50.7%. The percentages for H1 and H6 were 48.8% and 49.6%, respectively, which were only 0.2% higher than the expected values. Consequently, only H3 was deleted. Eight individuals were omitted from the analysis because four individuals exhibited maximum extreme scores and four individuals exhibited minimum extreme scores. An inspection of unidimensionality was then carried out using PCAR (see Table 5), demonstrating that the items fulfilled the criteria for unidimensionality. The four items explained a total of 65.00% of the variance, which is considered high.

Conduciveness (COND) Dimension

COND refers to the home-learning environment created by parents for their children. The infit and outfit ZSTD SD (2.7 and 2.6, respectively) for this item indicate misfit of items. CD5 and CD4 exhibit the highest violation of the infit and outfit ZSTD > 2.0. Hence, two items (CD5 and CD4) were deleted for this construct.

The fit of the selected items was then re-evaluated. The infit and outfit for the factor were well within the range. The infit ZSTD SD for this item improved tremendously from 2.7 for the original seven items subscale to 1.2 for the final four items subscale. Unidimensionality measures explained 48.70% of the factor, and the unexplained variance in the first contrast was 1.4, which was far below the cut-off value of 3.0.

Religiosity (REL) Dimension

Universal items concerning religious practices were posed to the respondents. The ZSTD SD, which is equivalent to 2.0 for item infit, shows that item redundancy occurs (see Table 3). The mean measure for person stands at 0.86, indicating overall easy-to-endorse items, and the reliability of the factor is 0.85. Item separation (SEPAR) is 7.36, indicating that the items were sufficiently spread out over a continuum. The item misfit order was then examined (see Table 4). Items that were at the top of the misfit order were R5, R2 and R6 in descending order. The infit standardised score for item R5 was 2.4, and the outfit standardised value for item R2 was 2.6. R6 seemed to be in a better position to be retained; however, the observed values of R4 were below the expected match value (see Table 4). Therefore, R4 was also deleted. The items deleted in this phase were R5, R2 and R4. Finally, the infit and outfit indices were re-evaluated for the items deleted subscales. The infit and outfit MNSQ for items were within the stipulated range of 0.7–1.4. The separation of items also increased, showing a wider spread of continuum being covered. Redundant and duplicated items measuring the same aspect in the latent trait were eliminated. The unidimensionality assumption was tested. The variance explained constituted 73.40%, and the unexplained variance in the first contrast was 1.5 (< 3.0) (see Table 5).

Control (CONT) Dimension

CONT refers to parental control and contains the most items in the instrument, initially consisting of 12 items. From the analysis, misfit order indicated that items C7, C6, C12 and C4 violated the acceptable ZSTD infit and outfit values of < 2.0. Hence, the final items retained were C1, C2, C3, C5, C8, C9, C10 and C11. After deletion, the analysis revealed that both the infit and outfit MNSQ = 1.01. PCAR analysis indicated that unidimensionality with unexplained variance in the first contrast improved from 2.0 to 1.9, and the variance explained by the measures was 56.80% (see Table 5).

Motivation (MOV) Dimension

The standardised infit value is –0.3 indicating overall overfitting. Hence, item deletion is conducted to trim the pool of items for this dimension. Next, the misfit order is generated (see Table 4). All items' infit and outfit MNSQ are within the range, except for M3 and M5. The measure of M6, which was –1.16, indicated that the item was difficult to endorse. Hence, it was also considered for item deletion. The items M5, M3 and M6 were deleted. The unexplained variance in the first contrast is 1.9, which is less than 3.0, indicating unidimensionality (see Table 5).

Warmness (WARM) Dimension

The warmness (WARM) dimension consisted of eight items. The infit standardised MNSQ was –0.1, which indicated slight overfit for this construct (see Table 3). Item reduction is also expedited. The misfit order shows that W8 and W5 (see Table 4) were probable candidates for elimination. The two items showed lower than expected match values (see Table 4). Therefore, only two items were selected for deletion (W8 and W5). Next, the unidimensionality measure is evaluated for the newly compiled items in the dimension. The eigenvalue of the raw variance explained is 9.3, which constituted 60.70% of the variance explained; the unexplained variance in the first contrast is 1.8, indicating unidimensionality (refer Table 5).

Conflict (CONF) Dimension

The CONF dimension contains six items. The misfit table generated was scrutinised, and X6 topped the list for item misfit. Furthermore, its standardised infit and outfit MNSQ > = 2.0, indicating item misfit (see Table 5). The point biserial correlation measure was also low (0.48). Item X2 also showed a lower than expected match value (Table 4). Consequently, X6 and X2 were deleted. The eigenvalue of the raw variance explained is 4.8, which constituted 54.30% of the variance explained; the unexplained variance in the first contrast is 1.5, indicating unidimensionality (see Table 5).

Table 3. Selected Model Fit information and separation measures of items using Rasch

Constr	Analysis phase	Item Infit			Item Outfit			SEPAR	RELIA
		MNSQ	ZSTD	ZSTD SD	MNSQ	ZSTD	ZSTD SD		
ASP	Initial	1.00	–0.1	3.2	0.99	–0.2	2.1	9.68	0.99
	Final	0.99	–0.1	0.7	0.97	–0.3	0.8	6.08	0.97
HWK	Initial	0.99	–0.5	2.9	1.01	–0.4	3.3	10.24	0.99
	Final	0.98	–0.4	1.2	0.96	–0.5	1.3	11.10	0.99
COND	Initial	1.01	0.1	2.7	1.02	0.1	2.6	4.37	0.95
	Final	0.99	–0.1	1.2	0.99	–0.5	1.1	3.14	0.91
REL	Initial	0.99	–0.2	2	0.97	–0.4	1.9	7.36	0.98
	Final	0.99	–0.1	1.0	0.96	–0.4	0.10	8.96	0.99
CONT	Initial	1.01	0.0	2.6	1.04	0.3	2.7	8.14	0.99
	Final	1.01	–0.1	2.3	1.01	–0.1	2.0	10.01	0.99
MOV	Initial	1.00	–0.3	3.0	1.01	0.0	3.3	9.77	0.99
	Final	0.99	–0.2	1.5	0.99	–0.1	1.8	3.94	0.94
WARM	Initial	1.00	–0.1	1.4	1.00	–0.1	1.5	10.1	0.99
	Final	1.00	0.0	0.7	1.00	0.0	0.8	11.87	0.99
CONF	Initial	1.00	0.0	1.2	1.01	0.1	1.4	10.4	0.99
	Final	0.99	–0.2	0.8	0.99	–0.1	0.8	11.23	0.99

Constr=Construct, MNSQ=mean square value, ZSTD=standardised score of MNSQ,
ZSTD SD=standardised score of standard deviation, SEPAR=separation index, RELIA=reliability.

Table 4. The infit and outfit measure of item misfit order

Constr	Item	MNSQ			Pt-measure	Exact	Match
		Infit	ZSTD	Outfit	(r_{bis})	Obs %	Exp %
ASP	A2	1.21		1.20	0.62	55.6	56.1
	A7	1.16		1.16	0.70	45.5	47.0
	A6	1.09		1.07	0.62	61.5	62.8
	A3	1.01		1.01	0.74	56.4	47.8
	A5	0.91		0.89	0.77	52.8	47.0
	A4	0.88		0.87	0.77	52.6	46.7
	A1	0.76		0.71	0.82	56.1	46.0
HWK	H3	1.62		1.82	0.48	40.7	50.7
	H1	0.96		0.93	0.67	48.8	48.6
	H6	0.92		0.94	0.69	49.6	49.4
	H7	0.90		0.93	0.68	53.1	52.9
	H2	0.89		0.83	0.76	48.5	44.7
	H4	0.84		0.82	0.74	45.6	45.3
	H5	0.80		0.78	0.78	49.1	44.6
COND	CD5	1.41		1.49	0.46	36.2	38.3
	CD4	1.19		1.20	0.55	32.7	37.6
	CD7	1.04		1.05	0.56	43.3	43.0
	CD2	0.94		0.92	0.56	51.2	48.2
	CD3	0.84		0.86	0.62	54.5	45.2
	CD6	0.85		0.80	0.61	53.1	46.1
	CD1	0.80		0.80	0.63	49.6	43.6
REL	R5	1.21		1.23	0.78	51.5	51.5
	R2	1.16		1.22	0.72	51.4	55.8
	R6	1.02		0.89	0.75	57.3	56.7
	R7	0.98		0.98	0.80	51.7	49.8
	R4	0.98		0.97	0.74	56.7	56.8
	R3	0.88		0.87	0.83	57.0	50.3
	R1	0.71		0.63	0.83	62.3	53.2

(continued on next page)

Table 4. *(continued)*

Constr	Item	MNSQ			Pt-measure	Exact	Match
		Infit	ZSTD	Outfit	(r_{bis})	Obs %	Exp %
CONT	C7	1.29		1.45	0.53	34.5	38.3
	C6	1.30		1.33	0.52	41	42.8
	C12	1.29		1.32	0.52	40.2	43.4
	C4	1.01		1.20	0.62	36.7	38.1
	C1	1.07		1.03	0.62	41.6	38.3
	C8	1.01		0.96	0.62	51.5	46.5
	C2	1.00		0.95	0.60	51.2	49.7
	C3	0.93		0.95	0.66	44.2	39.8
	C5	0.85		0.91	0.69	44.0	41.1
	C10	0.80		0.83	0.72	46.6	41.8
	C9	0.82		0.82	0.70	46.6	42.9
	C11	0.76		0.76	0.72	50.7	44.8
MOV	M5	1.41		1.51	0.56	43.6	46.3
	M3	1.25		1.34	0.65	40.4	42.6
	M6	1.11		1.09	0.57	56.4	58.4
	M7	0.94		0.94	073	47.2	43.7
	M4	0.86		0.88	0.71	50.9	45.8
	M8	0.82		0.82	0.74	52.3	46.4
	M1	0.80		0.77	0.75	55.3	46.4
	M2	0.77		0.73	0.77	56.6	46.8
WARM	W8	1.25		1.27	0.59	47.6	51.2
	W5	1.02		1.08	0.63	52.4	52.9
	W6	1.01		0.94	0.61	65.1	62.9
	W2	0.97		0.99	0.69	48.7	47.4
	W3	0.97		0.98	0.70	48.4	46.8
	W7	091		0.92	0.72	55.4	48.0
	W1	0.92		0.90	0.72	48.7	45.2
	W4	0.91		0.90	0.74	45.2	43.8

(continued on next page)

Table 4. *(continued)*

Constr	Item	MNSQ			Pt-measure	Exact	Match
		Infit	ZSTD	Outfit	(r_{bis})	Obs %	Exp %
CONF	X6	1.14		1.17	0.48	34.9	37.8
	X2	1.05		1.08	0.56	30.4	33.0
	X4	1.04		1.03	0.53	44.9	41.8
	X3	0.95		0.97	0.52	52.3	50.4
	X1	0.92		0.92	0.61	37.6	37.1
	X5	0.89		0.88	0.61	42.7	41.5

Table 5. First Contrast from PCAR

Constr	Raw variance explained		Unexplained variance in 1st contrast	
	Eigenvalue	Empirical	Eigenvalue	Empirical
ASP	7.4	65.10	1.5	12.70
HWK	11.1	65.00	1.9	32.40
COND	4.7	48.70	1.4	13.90
REL	11.0	73.40	1.5	10.00
CONT	10.5	56.80	1.9	10.20
MOV	7.3	59.40	1.9	15.10
WARM	9.3	60.70	1.8	11.80
CONF	4.8	54.30	1.5	16.90

DISCUSSION

Overall, the items were well endorsed by the respondents, implying that most items functioned well according to the constructs measured. They were within the acceptable range 0.70 and 1.30. However, the item ZSTD SD indicated item misfit for the overall measured items. The items indicated good fit on average. A look at the item separation index indicates that the items may be divided into eight categories (SEPAR for item = 8.31). The analysis revealed infit ZSTD SD = 2.1 for person, indicating that the responses were not consistent. PCAR results for the overall items showed that SPA was not unidimensional. This finding indicated that parental actions were most likely multidimensional. However, further confirmation may be necessary using multidimensional item response theory (MIRT).

The eight individual dimensions with 62 items exhibited good item fit with an initial MNSQ infit and outfit ranging from 0.97 to 1.04. After item deletion, the 43 items selected using Rasch yielded a range for MNSQ infit and outfit statistics of between 0.96 and 1.01. The initial standard deviation of the standardised infit (ZSTD SD) was beyond the cut-off value of 2.00 for all dimensions except WARM and CONF. This finding indicated both item misfit for the related dimensions and the need to re-examine some items. The WARM and CONF dimensions exhibited ZSTD SD values of 1.4 and 1.2, respectively, indicating no item misfit. However, the values observed for several items were lower than the expected values; therefore, these items were eliminated. Nineteen items were deleted from the instrument, and the ZSTD SD values improved with values below 2.0.

Negative values of the mean of the standardised infit and outfit (ZSTD) indicates overfit in the ASP construct, which suggests that the data fit the model better than expected and that redundancy may exist. Redundancy indicates the need to trim items to reduce the length of the instrument. This further supported item deletion in five dimensions, namely ASP, HWK, REL, MOV and WARM.

To test for unidimensionality within each individual construct, the items in each of the eight dimensions were tested for unexplained variance in the first contrast from the principal component analysis of standardised residuals. All dimensions exhibited eigenvalues of less than 3.0, indicating that the items are intact and measure what they should, with explained variances ranging from 48.70% to 73.40%.

CONCLUSION

Using Rasch approach, SPA demonstrated acceptable psychometric properties. With reference to SPA, selection using Rasch has a slightly different set of items as compared to SEM (refer to Table 6). However, the difference is not substantially large. The original SPA consisted of 34 items, whereas Rasch contained 43 items. Thirty-one items are similar in the SEM-based SPA. Approximately 91% of the items in the original SPA were selected using Rasch. The high consistency of item selection in SPA led to a preliminary conclusion that it is permissible to develop instruments using either SEM or Rasch theory, as the outcome will probably not be significantly different. Furthermore, if the content of the instrument does not differ much, it may support the accuracy and validity of the developed instrument. Analysis clearly demonstrated that parents' actions towards their children were of a multidimensional nature. In short, this study corroborated usability of and confidence in SPA. Further investigations and replications are necessary to improve the instrument.

Table 6. Comparisons of selected items using SEM and RASCH

Construct	Item	SEM	RASCH	Construct	Item	SEM	RASCH
ASP	A1	√	√	CONT	C1	√	√
	A2				C2		√
	A3	√	√		C3		√
	A4	√	√		C4		
	A5	√	√		C5		√
	A6				C6		
	A7				C7		
HWK	H1	√	√		C8		√
	H2	√	√		C9	√	√
	H3				C10	√	√
	H4	√	√		C11	√	√
	H5	√	√		C12	√	
	H6		√	MOV	M1	√	√
	H7		√		M2	√	√
COND	CD1	√	√		M3		
	CD2	√	√		M4	√	√
	CD3	√	√		M5		
	CD4				M6		
	CD5				M7		√
	CD6	√	√		M8	√	√
	CD7		√	WARM	W1	√	√
REL	R1	√	√		W2	√	√
	R2				W3	√	√
	R3	√	√		W4	√	√
	R4				W5		
	R5				W6		√
	R6	√	√		W7		√
	R7	√	√		W8		
				CONF	X1	√	√
					X2	√	
					X3	√	√
					X4		√
					X5	√	√
					X6	√	

ACKNOWLEDGEMENT

This study was funded by Institute of Postgraduate Studies Research Grant (308/AIPS415401), Universiti Sains Malaysia.

BIBLIOGRAPHY

Allen, L., & Oshagan, H. (1995). The UCLA loneliness scale: Invariance of social structural characteristics. *Personality and Individual Differences, 19*(2), 185–195.

Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika, 43*, 561–573.

Bartholomew, D. J., & Knott, M. (1999). *Latent variable models and factor analysis*. London: Arnold.

Bode, R. K., & Wright, B. D. (1999). Rasch measurement in higher education. In J. C. Smart, & W. G. Tierney (Eds.), *Higher education: Handbook on theory and research* Vol. XIV (pp. 287–316). New York: Agathon Press.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Bollen, K. A. (2002). Latent variables in psychology and social sciences. *Annual Review of Psychology, 53*, 605–634.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Embretson, S. E., & Reisse, S. P. (2000). *Item respose theory for psychologists*. Mahwah, NJ: Erlbaum.

Glockner-Rist, A., & Hoijtink, H. (2003). The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. *Structural Equation Modeling, 10*(4), 544–565.

Green, K. E., & Frantom, C. G. (2002). *Survey development and validation with the Rasch model.* Paper presented at the International Conference on Questionnaire Development, Evaluation and Testing, Charleston, SC, 14–17 November.

Griffin, P. (2005). Developing a masure of wealth for primary student families in a developing country: Comparison of two methods of psychometric calibration. *Studies in Educational Evaluation, 31*, 192–206.

Karabatsos, G. A. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement, 1*, 152–176.

Koh, S. Y. W., & Ong, S. L. (2010). Multidimensional parents' action on adolescents academic achievement in Malaysia. *Journal of International Education Studies, 3*(3), 179–191.

Lamoureux, E., Pallant, J., Pesudovs, K., Hassel, J., & Keeffe, J. (2006). The impact of vision impairment questionnaire: An evaluation of its measurement properties using Rasch analysis. *Invest Opthamol, 47*, 4732–4741.

Linacre, J. M. (2006). *A user's guide to Winsteps: Rasch model computer programs*. Chicago: John M. Linacre.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modelling in psychological research. *Annual Review of Psychology, 51*, 201–226.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Merrell, C., & Tymms, P. (2005). Rasch analysis of inattentive, hyperactive and impulsive behaviour in young children and the link with academic achievement. *Journal of Applied Measurement, 6*(1), 1–18.

Nijsten, T., Unaeze, J., & Stern, R. S. (2006). Refinement and reduction of the impact of psoriasis questionnaire: Classical test theory vs. Rasch analysis. *British Journal of Dermatology, 154*(4), 692–700.

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the hospital anxiety and depression scale (HADS). *British Journal of Clinical Psychology, 46*(1), 1–18.

Prieto, L., Alonso, J., & Lamarca, R. (2003). Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes, 1*, 1–27.

Smith, R. M., Schumacker, R. E., & Bush, M. J. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement, 2*(1), 66–78.

Waugh, R. F., & Chapman, E. S. (2005). An analysis of dimensionality using factor analysis (true-score theory) and Rasch measurement: What is the difference? Which method is better? *Journal of Applied Measurement, 6*(1), 80–99.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know* (pp. 65–104). Mahwah, NJ: Lawrence Erlbaum.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago: Mesa press.