

ORIGINAL REPORT

RELIABILITY AND VALIDITY OF THE PAEDIATRIC VOLITIONAL QUESTIONNAIRE – CHINESE VERSION

Li-Ting Liu, MSc, OT Certified (Taiwan)¹, Ay-Woan Pan, PhD, OTR/OTC (USA, Taiwan)^{1,2}, Lynn Chung, PhD³, Susan Shur-Fen Gau, MD, PhD^{2,4}, Jessica Kramer, PhD, OTR/L (USA)⁵ and Jin-Shei Lai, PhD, OTR (USA)⁶

From the ¹School of Occupational Therapy, College of Medicine, National Taiwan University, ²Department of Psychiatry, National Taiwan University Hospital, ³Department of Statistics, National Taipei University, ⁴Department of Psychology, Graduate Institute of Brain and Mind Sciences, Graduate Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan, ⁵Department of Occupational Therapy, Boston University, Boston, USA and ⁶Department of Medical Social Science, Feinberg School of Medicine, Northwestern University, Chicago, USA

Objective: The purpose of this study is to examine the psychometric properties of the Chinese version of the Paediatric Volitional Questionnaire (PVQ-C) for use amongst preschoolers in Taiwan.

Methods: Forty preschoolers with developmental delays were randomly selected from northern Taiwan, along with another 40 typically developing preschoolers. The data was analysed using Rasch measurement model for construct validity and classical test theory for item reliability, intra- and inter-rater reliability, and convergent validity.

Results: The results indicated the PVQ items of PVQ-C fit into a unidimensional continuum of volition (logit –6.63–3.05) with 2 items representing differential item functioning for diagnostic group. The person-fit statistics showed that 83% participants' response could be appropriately estimated and stratified (separation index=1.86). Using a modified 3-point rating scale resulted in acceptable item reliability (0.97), intra-rater reliability (0.412–1.0), 86% PVQ items in test-retest reliability >0.4, and convergent validity ($r=0.562-0.656$).

Conclusions: The PVQ-C is regarded as a reliable and valid instrument for assessing the volitional status of children, as a reference for subsequent clinical judgement, and for client-centred intervention programmes and treatment effects.

Key words: evaluation; motivation; developmental delays; psychometric properties; Rasch model.

J Rehabil Med 2013; 45: 99–104

Guarantor address: Ay-Woan Pan, School of Occupational Therapy, College of Medicine, National Taiwan University. Correspondence address: Room 407, No. 17 Xu-Zhou Road, Taipei 100, Taiwan. E-mail: aywoan@ntu.edu.tw

Submitted November 22, 2011; accepted June 19, 2012

INTRODUCTION

Researchers propose that the rehabilitative benefit of everyday activities lies in the inherent social value, meaningfulness and intrinsically motivating nature of such activities (1–3). Motivation is widely recognized as one of the many factors affecting goal attainment and treatment outcomes in rehabilitation (4), but literature relating to the assessment of child motivation

within the context of rehabilitation is extremely limited. Furthermore, in published attempts to gain an understanding of child motivation, the use of assessment in rehabilitation is invariably reliant upon the verbal and cognitive ability of the children as the means of identifying both their interests and their articulated feelings; thus, such approaches have limited application for young children or those with cognitive impairment (5).

It is, nevertheless, the case that these children may derive the greatest benefits from interventions specifically designed to appeal to their interests and values. It should also be noted that most examples of such assessment tend to be guided by an underlying theoretical assumption that motivation is static and inherent (1, 6–12). Therefore, the motivation for task engagement may be context dependent and constantly shifting according to the meaningfulness and value, as well as the extent to which the environment itself provides adequate support. These assessments do not consider sufficiently how these factors affect a child's performance, especially for disabled children. Nowadays, the assessment of motivation for use in Taiwan is severely limited. There is a clear requirement for alternative methods of assessing and understanding the motivation of children to engage in everyday activities so as to ensure best practice in paediatric rehabilitation in Chinese-speaking countries.

The Paediatric Volitional Questionnaire (PVQ) (5, 13, 14) was developed as a means of assessing and explaining the motivation of children towards particular everyday activities. It examines specific types of behaviour regarded as being indicative of motivation, as well as the environmental factors. Therefore, it is also appropriate for use with young children and those who are unable to either express their motivation to engage in specific activities or to articulate their feelings about specific everyday occupations. Furthermore, this instrument can be used to plan effective client-centred rehabilitation programmes.

The PVQ is based upon the concepts of the Model of Human Occupation (MOHO) (15), a model that uses the concept of volition to gain an understanding of the types of behaviour related to motivation and engagement in everyday activities.

Volition refers to the dynamic interactions that occur between the personal causation (e.g. the perspectives that exist amongst individuals with regard to their capacity and efficacy to engage in everyday activities) (15), values (e.g. essentially the perceptions amongst individuals about what is “good”, “right” and “important”, which influence the choices and preferences amongst children to engage in certain activities over others), and interests of an individual to engage in everyday activities, and reflects a situation that is subject to change, depending upon the everyday activity involved and the context in which it takes place (15). People will invariably take an interest in activities that they prefer to do or that they find enjoyable. For example, when children find video games funny and enjoyable, they are internally rewarded and motivated to continue to engage in such games, even when the games become more complex or challenging (16). The MOHO also recognizes that both the social environment (including the attitudes of individuals and the accepted ways in which activities are performed) and the physical environment (including both space and objects) can support or constrain engagement in activities (17). The extant literature provides general support for the clinical application of the MOHO in paediatric populations (5, 13, 16, 18–20).

All of these factors influence the willingness amongst children to engage in activities in the context of rehabilitation, as well as everyday life; however, it must also be recognized that they are responsive to differences between cultures (17, 18). The research aims of the present study are that the Chinese version of the PVQ (PVQ-C) has acceptable construct validity, convergent validity, intra-rater and test-retest reliability. The researchers translated the English version of the PVQ (2nd edition) (18) into Chinese using a two-step procedure (21, 22), with both the readability and consistency of the translated version being extensively checked by means of several pilot studies (23).

METHODS

The research protocol was approved by the Research Ethics Committee of the National Taiwan University Hospital (IRB NO. 9361700379). All participants and their parents provided written informed consent.

Participants

At the beginning of the study, we investigated the reachable registered nursery schools, kindergartens, special education centres and hospitals with early intervention centres in Northern Taiwan. Qualified centres were randomly selected and letters were sent to invite them to participate. For the developmental delays (DD) participants, the inclusion criteria were: (i) formal diagnosis of developmental delays by attending paediatricians or psychiatrists; (ii) age between 3 and 6 years; and (iii) the participants consent to participate in the study. The teachers or the therapists assisted to recruit the participants. For the typically developing (TD) participants, the inclusion criteria were: (i) confirmed normal development by “the Zero-To-Six-Year-Old Child Development Screening Checklist”, (ii) age group matching with the developmental delay subject group; and (iii) participants consent to participate in the study. The exclusion criterion for TD participants was a history of neurological or psychological/psychiatric condition.

Both the purpose of this study and the procedure were fully explained to the parents of the participants, with written permission being secured prior to proceed with the data collection. A final group of TD preschoolers ($n=40$) and a group of preschoolers with DD ($n=40$) were recruited to the study (Table I).

Measures

Zero-to-six-year-old child development screening checklist. The “zero-to-six-year-old child development screening checklist” (24) was used to confirm the developmental status of the participants. The screening checklist is designed to be used with PVQ items under Taiwanese culture for early detection of developmental delays from newborns up to children aged 71 months, and has been found to have both acceptable concurrent validity and predictive validity (24).

Paediatric Volitional Questionnaire-Chinese version (PVQ-C). The PVQ-C was translated by standard translation protocol based on previous literature (21, 22). The original PVQ was translated by a senior occupational therapist (10 years of working experience) into Chinese, and then was back-translated into English by a bilingual occupational therapist. These parallel versions were reviewed by our translation committee to ensure the congruency of meaning between original and back-translated versions. There were 3 sentences showing questionable translation results, which were thus re-translated and back-translated. The final version was examined by a committee comprised of 5 professionals (i.e. 1 rehabilitation physician, 1 special education teacher, 1 paediatric physical therapist, and 2 occupational therapists), resulting in a face validity of 0.87. This version was used as a pre-test version in a pilot study for feasibility testing in clinical practice. The invited raters found that there were no difficulties in understanding the PVQ-C instruction and the definition of PVQ items as well as their behavioural characteristics. The inter-rater reliability of the pilot test reached 0.7. The final version of PVQ-C comprises 14 PVQ items representing volitional behaviour indicative of personal causation, values and interests for the children aged between 2 and 7 years. These PVQ items are rated using a 4-point Likert-type scale (“passive”, “hesitant”, “involved” and “spontaneous”), each of which indicates the amount of support that the children require in order to demonstrate each type of behaviour. The PVQ items are designed to reflect a theoretical range of types of volitional behaviour, from more simple to more complex.

Table I. Demographic characteristics of the participants

Variable	Normal group ($n=40$)	DD group ($n=40$)	Total ($n=80$)	χ^2	p
Gender, male, n (%)	17 (42.5)	27 (67.5)	44 (55)	5.05	0.03*
Age, n (%)					
24–35 months	3 (7.5)	3 (7.5)	6 (7.5)		
36–47 months	10 (25.0)	10 (25.0)	20 (25.0)		
48–59 months	11 (27.5)	11 (27.5)	22 (27.5)		
60–71 months	11 (27.5)	11 (27.5)	22 (27.5)		
72–83 months	5 (12.5)	5 (12.5)	10 (12.5)		
Total, mean (SD)	55.9 (14.7)	63 (10.9)	58 (14.2)	–	–
Education, n (%)					
Father				16.40	0.01*
<6 years	–	2 (5.0)	2 (2.5)		
6–9 years	–	4 (10.0)	4 (5.0)		
9–14 years	5 (12.5)	14 (35.0)	19 (23.8)		
>12 years	24 (60.0)	11 (27.5)	35 (43.8)		
Mother				17.32	0.00**
<6 years	–	1 (2.5)	1 (1.3)		
6–9 years	1 (2.5)	3 (7.5)	4 (5.0)		
9–14 years	6 (15.0)	16 (40.0)	22 (27.5)		
>12 years	25 (62.5)	8 (20.0)	33 (41.3)		
Family income				27.35	0.00**
<20,000 NT\$	–	4 (10.0)	4 (5.0)		
20,000–88,000 NT\$	22 (55.0)	29 (72.5)	51 (63.9)		
>88,001 NT\$	18 (45.0)	6 (15.0)	24 (30.0)		
Missing values	0	1	1		

* $p<0.05$; ** $p<0.01$.

SD: standard deviation; NT\$: New Taiwan dollars.

The PVQ also includes an extensive checklist used to document aspects of the physical and social environment that may have direct (or indirect) impacts on volition. In carrying out the PVQ rating, the examiner observed each child engaging in one or two different everyday activities for 15–30 min in various contexts (such as a classroom, therapeutic space, play room, playground or home), with these PVQ items having been specifically designed to measure a theoretical range of “less complex” to “more complex” volitional behaviour (5). Several earlier studies on the English version of the PVQ provide support for both the construct validity (item reliability=0.8–0.95) (5, 14, 19) and inter-rater reliability (rater reliability=0.95) (5, 19) of the instrument.

Procedures

The primary care-giver completed the demographic data sheet, after which the participants were observed and videotaped in two situations. The first was in an educational or therapeutic situation; for example, in a classroom or during a treatment session. The participants themselves were encouraged to select the second situation (if any of the participants were unable to decide, then the caregivers would choose activities that they believed were their favourites).

After watching the videotapes, the researcher rated each participant using the PVQ-C. A total of 20 of the participants were selected for “test-retest reliability” studies, and were subsequently tested again 2 weeks later. A further 20 videotapes were selected for “intra-rater reliability”, with the rater being unaware of the developmental status of the participants.

Data analysis

The Rasch measurement model (25) was adopted for the analysis of the PVQ ratings using the Facets for Windows program (Version 3.62, MESA Press, Chicago, IL, USA) (26, 27). The Rasch ratings scale model converts the ordinal data into interval measures, calibrates the PVQ items and participants onto an underlying measurement continuum (i.e. volition), and then goes on to evaluate the fitness of the PVQ items (28, 29). The criteria used in the present study were applied to judge the fit and adequacy of the test included (26, 27, 30, 31):

- *item calibration*, representing item difficulty, or the position of the item along the underlying continuum, which will be used to investigate if the level of difficulty of the PVQ items are adequate to discriminate the level of volitional status for the participants and if more items are needed;
- *fit statistics*, which indicate the extent to which the data is well fitted to the Rasch measurement model for measuring the underlying continuum. The criteria of fit in this study was set at an Infit mean

square (MnSq) of between 0.6 and 1.4, with an associated standardised z-score (ZSTD) of between 2 and -2 (26, 30);

- *item separation index*, which indicates the way in which the PVQ items are spread along the underlying continuum to measure the less to more complex types of behaviour (26, 30) which represents if the PVQ items are sensitive enough to identify different level of volitional status of the participants;
- *item separation reliability*, indicating the reliability of the separation of the PVQ items along the underlying continuum. A level in excess of 0.9 was regarded as being acceptable (32) and;
- *differential item functioning* (DIF), which indicates whether the test functions across different age and gender groups. A Mantel-Haenszel procedure was used to compare whether the item calibration differed across different groups (30). Assessment tools are often found to be client-specific, and must therefore be closely examined for item bias prior to being applied within clinical practice (27, 33).

The data on all of the participants were used to examine the fit of the PVQ items with the measurement model, as well as the item validity and internal consistency, with SPSS 11.0 (SPSS Inc., Chicago, IL, USA) being used for the descriptive analyses, and test-retest and intra-rater reliability being analysed by weighted Kappa (34, 35). Finally, the Spearman's rank correlation coefficients were analysed between the PVQ-C scores and the zero-to-six-year-old child development screening checklist.

RESULTS

Paediatric Volitional Questionnaire – Chinese version construct validity

The results revealed that the rating scale of 2 (“hesitant”) was rarely used, ranging between 5% and 56%; consequently, rating 2 and rating 3 (“involved”) were combined into rating 2, with the original rating 4 then becoming rating 3. The converted scale was then analysed again, after which the new 3-point scale was used to carry out the following analyses (all items have ordered threshold).

The Infit MnSq and ZSTD of all PVQ items were found to have good fit with the model requirements (i.e. $0.6 < \text{infit} < 1.4$, $-2 < \text{ZSTD} < 2$), thereby providing clear confirmatory evidence that the PVQ items are measuring a uni-dimensional continuum of volition (Table II). The continuum of item difficulty was

Table II. *Item statistics*

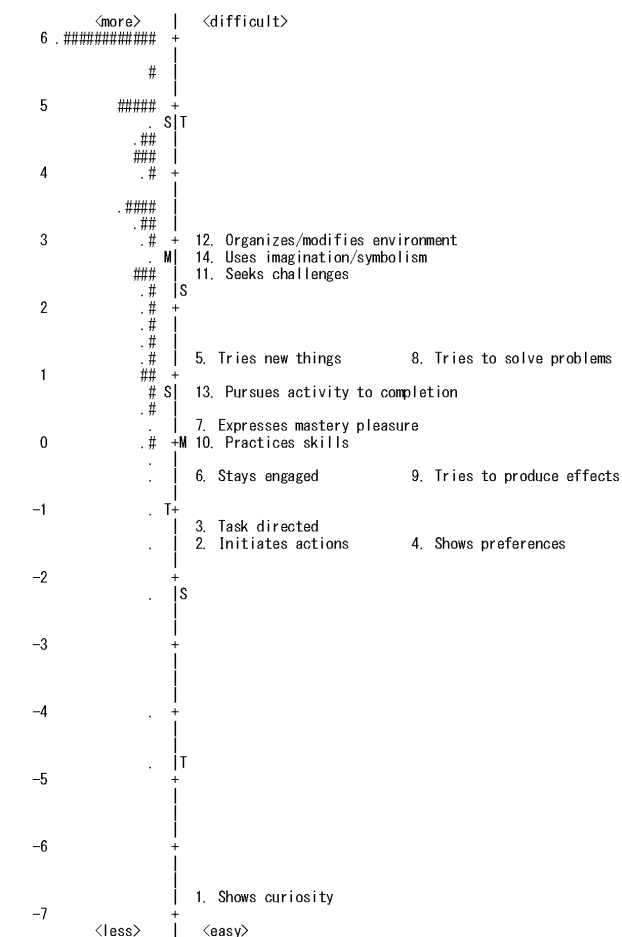
Item number	PVQ item	Measures (logits)	Error (logits)	Infit		Outfit	
				MnSq	ZSTD	MnSq	ZSTD
12.	Organizes/modifies environment	3.05	0.42	1.06	0.2	1.05	0.3
14.	Uses imagination/symbolism	2.77	0.20	1.20	1.4	1.15	0.8
11.	Seeks challenges	2.56	0.16	1.12	1.1	1.15	0.9
5.	Tries new things	1.18	0.20	0.93	-0.6	0.83	-0.9
8.	Tries to solve problems	1.13	0.17	0.99	-0.1	0.91	-0.6
13.	Pursues activity to completion	0.80	0.17	0.94	-0.5	1.01	0.1
7.	Expresses mastery pleasure	0.37	0.19	1.10	0.7	0.82	-0.9
10.	Practices skills	0.00	0.19	0.67	-2.8	0.70	-1.3
9.	Tries to produce effects	-0.45	0.27	1.30	1.5	1.58	1.4
6.	Stays engaged	-0.48	0.21	0.76	-1.8	0.53	-1.8
3.	Task directed	-1.35	0.25	0.73	-1.8	0.62	-0.8
2.	Initiates actions	-1.41	0.25	0.97	-0.2	2.37	2.2
4.	Shows preferences	-1.54	0.30	1.43	1.8	1.87	1.3
1.	Shows curiosity	-6.63	1.14	1.57	0.6	0.20	-1.7
Mean (SD)		0.00 (2.35)	0.30 (0.24)	1.05 (0.25)	0.0 (1.3)	1.06 (0.54)	-0.1 (1.2)

PVQ: Paediatric Volitional Questionnaire; SD: standard deviation. MnSq: mean square; ZSTD: standardised z-score

shown where simpler types of volitional behaviour had lower values on the continuum (i.e. “shows curiosity”) and complex types of volitional behaviour had higher values on the continuum (i.e. “organizes/modifies environment”). The order of item difficulty was found to be similar to that reported in the English studies (19).

There are 12% observations counted as misfit response. The person separation was 1.86 with a reliability of 0.77, who were distributed along the volition continuum from -4.76 to 6.85 (mean 1.06 (standard deviation (SD) 0.49)) with a positive skewed tendency (Fig. 1). There was 17% response yielding the ceiling score (e.g. spontaneous), 80% of which was made by the TD group, indicating that 83% response could be appropriately estimated by PVQ-C. Regarding the separation index, the PVQ-C divided our participants into 3 levels of volitional status.

The item separation index was 5.09, indicating that the PVQ items represented 8 levels of complexity along the underlying continuum (Table II). There were no DIF for gender, age level, occupational forms, or environments. However, there are two



Each “#” indicates 4

Fig. 1. Item-person map. The right side represents the construct of Paediatric Volitional Questionnaire, and the left side visualizes the person’s volitional status.

Table III. Intra-rater reliability (n = 20)

Item No.	PVQ item	Weighted kappa
1.	Shows curiosity	1.000
2.	Initiates actions	1.000
3.	Task directed	0.847
4.	Shows preferences	0.462 (n = 7)
5.	Tries new things	0.590
6.	Stays engaged	0.500
7.	Expresses mastery pleasure	0.583
8.	Tries to solve problems	0.900
9.	Tries to produce effects	0.588 (n = 7)
10.	Practices skills	0.727
11.	Seeks challenges	0.755
12.	Organizes/modifies environment	1.000 ^a
13.	Pursues activity to completion	0.886
14.	Uses imagination/symbolism	0.412 (n = 10)

^aData shown here is based on only one subject.

PVQ: Paediatric Volitional Questionnaire.

DIF PVQ items between DD and TD groups after conducting Bonferroni correction (*p* threshold set as 0.0357). They were item 9 – “Tries to produce effects (easier in DD group, *p* = 0.0008)”, and item 14 – “Uses imagination/symbolism (easier in DD group, *p* = 0.0006)”.

Convergent validity

The results of the Spearman’s correlation analysis revealed that PVQ-C was moderately correlated with the language/communication, social/personal, gross motor skill, fine motor skill and perception/cognition subtests (*r* = 0.562~0.656, *p* < 0.01).

Reliability

The Rasch item reliability on the 3-point scale was found to be 0.96, with the weighted Kappa for the randomly selected and re-scored PVQs being between 0.412 and 1.0, indicating acceptable intra-rater agreement (Table III).

Finally, the test-retest reliability was analysed using weighted Kappa, with the results showing that 86% of the PVQ items had kappa values in excess of 0.4, thereby indicating moderate to good consistency (Table IV).

Table IV. Test-retest reliability (n = 20)

Item No.	PVQ item	Weighted kappa
1.	Shows curiosity	1.000
2.	Initiates actions	0.704
3.	Task directed	0.659
4.	Shows preferences (n = 12)	0.520
5.	Tries new things (n = 20)	0.747
6.	Stays engaged	0.714
7.	Expresses mastery pleasure	0.557
8.	Tries to solve problems (n = 18)	0.737
9.	Tries to produce effects (n = 8) ^a	0.143
10.	Practices skills	0.708
11.	Seeks challenges (n = 20)	0.824
12.	Organizes/modifies environment (n = 7)	0.000
13.	Pursues activity to completion	0.569
14.	Uses imagination/symbolism (n = 8)	1.000

^aThe item shown here has low consistency.

DISCUSSION

This study is believed to be the first of its kind to examine the psychometric qualities of the Chinese version of the PVQ amongst Taiwanese participants, with our findings indicating that the use of a modified 3-point rating scale, within which the middle ratings of “hesitant” and “involved” are combined, provides better results.

This revised rating scale structure may be clinically more meaningful than the original 4-point structure, since clinicians have reported difficulty in determining the difference between the rating categories of “involved” (needs a little encouragement or mild assistance) and “hesitant” (needs considerable encouragement or assistance). Previous studies on the English PVQ reported similar findings (5, 19).

Overall, the evidence suggests that the PVQ-C has good construct validity. All 14 PVQ items are found to fit the Rasch measurement model, indicating that all of the PVQ items are valid measures of a volitional behaviour continuum. Furthermore, the distribution of those PVQ items along the continuum that are aligned with the theoretical assumptions underlying the assessment are found to be similar to the findings of the earlier studies. The results suggest that by observing the behaviour of children during task performance, therapists can obtain information on their volitional status in various natural contexts without specific cognitive ability or verbal expression. Furthermore, volitional status is found to be related to the developmental status of the children (e.g. with or without delays), thereby demonstrating that developmental challenges may have direct impacts on volitional status or vice versa. Since no DIF were discernible amongst age, gender, occupational forms, or environments, we do not need to adjust the item calibrations for different sub-groups. The professionals would better consider the characteristics of contexts to maximize clients’ participation, whereas MOHO emphasizes the importance of environment for facilitating and supporting the volitional function for individuals. In addition, items 9 and 14 need more caution when set as goals in intervention programmes for children with developmental delays, due to the discrepancy from TD children.

The results of this study provide strong support for the reliability of the PVQ, with the evidence suggesting that the PVQ can be used repeatedly, and across time, by the same rater. The positive change in PVQ score of individuals could be considered as the progress of treatment. Some of the previous studies have indicated that, for any rater who had been previously unfamiliar with a client, the greatest challenge was found to be the problem of rating the client for certain PVQ items, prime examples being “seeks challenges” and “tries new things” (5). These problems were also found to occur in the present study (“tries new things”); it therefore appears to be necessary for raters to become familiar with the child through direct reports from teachers or parents.

With regard to intra-rater reliability, raters were found to have problems deciding whether certain types of behaviour represented a particular score; thus, we suggest that the PVQ-C manual (36) be revised accordingly to include more detailed

information on the item definition and rating examples, in order to improve the consistency of the ratings. The manual should also indicate that those administering the PVQ-C should try to select a situation in which the child is certain to encounter a novel situation. To promote accurate ratings, therapists must read the PVQ-C manual carefully and be familiar with the theoretical concepts upon which the PVQ-C is based, and there may also be a need for therapists to attend training sessions or workshops prior to using the PVQ-C in practice, due to the complexity of the concepts measured by the instrument.

There are several limitations of the present study. Firstly, the sample size is too small and 2–3-year-old participants were less representative. Secondly, despite the high consistency of rating between raters in our pilot study, the raters’ severity cannot be examined in the study due to the single rater. Thirdly, there is no appropriate item for those participants who obtained the ceiling score, especially for those with developmental delays. This result suggests that more challenging PVQ items should be added in the future. All of these issues must be taken into consideration prior to any attempt to generalize the findings.

In conclusion, a firm understanding of client motivation is of significant importance in promoting skills development and functional performance amongst children in client-centred practice (37–39). Many practitioners are found to have difficulty evaluating the motivation of children with impaired cognitive capabilities or the ability to verbally express themselves. The results of this study reveal that the PVQ-C is a reliable and valid observational assessment tool for use in the assessment of motivation amongst children to engage in everyday activities in their own naturalistic contexts. Following the Rasch transformation of the raw scores into linear logit scores, we can more effectively plan intervention programmes for individuals based on their volitional status with no developmental frame, facilitate the higher order of volition through appropriate modalities, and encourage the client to engage in activities of their own free will.

We suggest that the “hesitant” and “involved” ratings should be merged to form a 3-point rating scale, since this may provide a better structure and enhanced informative performance. Although participants in this study were selected randomly, we expect that the recruitment of sources of participants in wider regions may help to increase the representative quality of the sample.

In order further to examine the reliability and validity of the PVQ, other suggested areas for future research include inter-rater reliability of different professional backgrounds, and concurrent validity with other motivation-related child assessments, such as the Motivation Assessment Scale (MAS) (40) or the Nowicki-Strickland Locus of Control Scale for Children (N-S scale) (41) and the Dimension of Motivation Questionnaire (42).

ACKNOWLEDGEMENT

The authors would like to thank the participants of the study for their contribution. This study was supported by grants from the Department of Health at the Executive Yuan, Taiwan (Grant No. DOH 95-TD-M-113-047) and from the National Science Council at the Executive Yuan, Taiwan (grant no. NSC 97-2628-S-002-001-MY3).

REFERENCES

- Bober SJ, Humphry R, Carswell HW, Core AJ. Toddlers' persistence in the emerging occupations of function play and self-feeding. *Am J Occup Ther* 2001; 55: 369–376.
- Hasselkus BR. 2006 Eleanor Clarke Slagle Lecture—the world of everyday occupation: real people, real lives. *Am J Occup Ther* 2006; 60: 627–640.
- Wood W. It is jump time for occupational therapy. *Am J Occup Ther* 1998; 52: 403–411.
- Sharrott GW, Cooper-Fraps C. Theories of motivation in occupational therapy: an overview. *Am J Occup Ther* 1986; 40: 249–257.
- Geist R. The validity study of the pediatric volitional questionnaire. Master's thesis. Chicago, IL: University of Illinois at Chicago; 1998.
- Cross L, Coster W. Symbolic play language during sensory integration treatment. *Am J Occup Ther* 1997; 51: 808–814.
- Bundy AC. Play and playfulness: what to look for. In: Parham LD, Fazio LS, editors. *Play in occupational therapy for children*. St Louis: Mosby; 1997. Chapter 4: 52–66.
- Dunkerley E, Tickle-Degnen L, Coster WJ. Therapist-child interaction in the middle minutes of sensory integration treatment. *Am J Occup Ther* 1997; 51: 799–805.
- Leipold EE, Bundy AC. Playfulness in children with attention deficit hyperactivity disorder. *Occup Ther J Res* 2000; 20: 61–82.
- Okimoto AM, Bundy A, Hanzlik J. Playfulness in children with and without disability: measurement and intervention. *Am J Occup Ther* 2000; 54: 73–82.
- Petri HL. Introduction. In: Petri HL, editor. *Motivation: theory, research, and applications*. 4th edn. CA: Brooks & Cole; 1996, p. 3–22.
- Pierce DE. Early object rule acquisition. *Am J Occup Ther* 1991; 45: 438–449.
- Chern J, Kielhofner G, de las Heras C, Magalhaes L. The volitional questionnaire: psychometric development and practical use. *Am J Occup Ther* 1996; 50: 516–525.
- Anderson S, Kielhofner G, Lai JS. An examination of the measurement properties of the pediatric volitional questionnaire. *Phys Occup Ther Pediatr* 2005; 25: 39–57.
- Kielhofner G. *Model of human occupation: theory and application*. 4th ed. Hagerstown: Lippincott Williams & Wilkins; 2007.
- Sholle-Martin S. Application of the Model of Human Occupation: assessment in child and adolescent psychiatry. *Occup Ther Ment Health* 1987; 7: 3–22.
- Reekmans M, Kielhofner G. Defining occupational therapy services in child psychiatry: an application of the model of human occupation. *Ergotherapie* 1998; 5: 6–13.
- Basu S, Kafkes A, Geist R, Kielhofner G. A user's guide to the Pediatric Volitional Questionnaire (version 2.0). Chicago, IL: University of Illinois at Chicago; 2002.
- Andersen SP. Using the pediatric volitional questionnaire to assess children with disabilities. Master's thesis. Chicago, IL: University of Illinois at Chicago; 1998.
- Kielhofner G, Mallinson T, Forsyth K, Lai JS. Psychometric properties of the second version of the Occupational Performance History Interview (OPHI-II). *Am J Occup Ther* 2000; 55: 260–267.
- Colon H, Haertlein C. Spanish translation of the Role Checklist. *Am J Occup Ther* 2002; 56: 586–589.
- Del Greco L, Walop W, Eastridge L. Questionnaire development: 3. Translation. *Can Med Assoc J* 1987; 136: 817–818.
- Liu LT. Psychometric Properties of the Chinese version of the Pediatric Volitional Questionnaire. Master's thesis. Taipei: National Taiwan University; 2004.
- Huang HL. Zero-to-six-year-old child development screening checklist. Taipei: Psychological Publishing Co.; 2000.
- Wright BD, Master GN. Rating scale analysis. Chicago, IL: MESA Press; 1982.
- Brown T, Unsworth C. Construct Validity of the test of visual-motor integration: an evaluation using Rasch analysis. *Hong Kong J Occup Ther* 2009; 19: 2–13.
- Farin E, Fleitz A. The development of an ICF-oriented, adaptive physician assessment instrument of mobility, self-care, and domestic life. *Int J Rehabil Res* 2009; 32: 98–107.
- Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil* 1989; 70: 857–860.
- Arnadottir G, Fisher AG. Rasch analysis of the ADL scale of the A-ONE. *Am J Occup Ther* 2008; 62: 51–60.
- Clemson L, Bundy A, Unsworth C, Singh MF. Validation of the modified assessment of living skills and resources, an IADL measure for older people. *Disabil Rehabil* 2009; 31: 359–369.
- Lai JS, Cella D, Kupst MJ, Holm S, Kelly ME, Bode RK, et al. Measuring fatigue for children with cancer: development and validation of the pediatric Functional Assessment of Chronic Illness Therapy-Fatigue (pedsFACT-F). *J Pediatr Hematol Oncol* 2007; 29: 471–479.
- Dreer LE, Berry J, Rivera P, Snow M, Elliott TR, Miller D, et al. Efficient assessment of social problem-solving abilities in medical and rehabilitation settings: a Rasch analysis of the Social Problem-Solving Inventory-Revised. *J Clin Psychol* 2009; 65: 653–669.
- Haley SM, Fragala-Pinkham MA, Dumas HM, Ni P, Gorton GE, Watson K, et al. Evaluation of an item bank for a computerized adaptive test of activity in children with cerebral palsy. *Phys Ther* 2009; 89: 589–600.
- Hinderer SR, Hinderer KA. Quantitative methods of evaluation. In: Joel A, DeLisa J, editors. *Rehabilitation medicine: principles and practice*. 2nd ed. Philadelphia: Lippincott Co.; 1993, p. 96–121.
- Cohen J. Weighted kappa – nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; 70: 213–220.
- Liu LT, Pan AW. A user's guide to the Pediatric Volitional Questionnaire – Chinese version. MOHO center in Taiwan; 2007.
- Yarrow L, McQuiston S, MacTruk R, McCarthy M, Klein R, Vietze P. Assessment of mastery motivation during the first year of life: contemporaneous and cross-age relationships. *Dev Psychol* 1983; 19: 159–171.
- Mizuno K, Tanaka M, Fukuda S, Imai-Matsumura K, Watanabe Y. Relationship between cognitive function and prevalence of decrease in intrinsic academic motivation in adolescents. *Behavioral and Brain Functions: BBF* 2011; 7: 4.
- Cordier R, Bundy A, Hocking C, Einfeld S. A model for play-based intervention for children with ADHD. *Austral Occup Ther J* 2009; 56: 332–340.
- Durand VM, Crimmins DB. Identifying the variables maintaining self-injurious behavior. *J Autism Dev Disord* 1988; 18: 99–117.
- Li HC, Lopez V. Chinese translation and validation of the Nowicki-Strickland Locus of Control Scale for Children. *Int J Nurs Stud* 2004; 41: 463–469.
- Sparks TA, Hunter SK, Backman TL, Morgan GA, Ross RG. Maternal parenting stress and mothers' reports of their infants' mastery motivation. *Infant Behav Dev* 2012; 35: 167–173.