Full Title

Convergent and divergent validity of the Participation and Environment Measure for Children and Youth (PEM-CY)

Short Title

Construct validity of the PEM-CY

Author Information

Fiona P Graham (PhD, B OccTher)

Senior Lecturer

University of Otago

34 Gloucester St, Christchurch 8013, New Zealand

William J Taylor MBChB, PhD, FRACP, FAFRM

Rehabilitation Teaching & Research Unit,

University of Otago Wellington

New Zealand

Cherie Le Lievre (MHSc, B Occ Ther)

Rehabilitation Teaching & Research Unit,

University of Otago Wellington

New Zealand

Contributions: Cherie Le Lievre undertook all aspects of this study within her masters' thesis studies under the supervision of Fiona Graham and Will Taylor. Fiona Graham completed the first draft of this manuscript with comments provided by Cherie Le Lievre and Will Taylor.

Word Count: 3733

Keywords: social participation, International Classification Functioning-Disability and Health (ICF); psychometrics, evaluation research, rehabilitation

ABSTRACT

Introduction: Robust measures of children's participation and environments are critical to understanding the impact of health and social conditions on children and young people. One widely used measure, the Participation and Environment Measure for Children and Youth (PEM-CY) has undergone preliminary psychometric analysis that appears favourable. This study extends these analyses and examines the construct validity of the PEM-CY through analysis of its convergent and divergent validity.

Methods: PEM-CY scoring was modified to enable comparison with other measures. A priori predictions of correlations between other measures of participation (Child and Adolescent Scale of Participation) and environments (Child and Adolescent Scale of Environment) and parental competence (Parenting Sense of Competence Scale) were examined using Spearman's rank correlation. Participants were university staff with children aged 5 to 17 years, recruited through an email invitation to participate in an online questionnaire.

Results: Based on responses from (*n*=67 to 100) parents, construct validity criteria were not met for the majority of PEM-CY summary scores, using either the recommended or modified scoring. Overall, >75% of predicted relationships were observed for some of the Environmental summary scores but for none of the Participation summary scores.

Conclusions: The convergent and divergent construct validity of summated scores from the PEM-CY was not confirmed. Findings reflect the inherent difficulty of measuring multi-dimensional constructs such as participation. Formative models of measurement design, developed using conjoint analysis of discrete choice data are suggested as a way to design more valid participation measures.

INTRODUCTION

Measurement of children's participation in activities and social roles (World Health Organisation, 2001) provides critical information about the impact of health and social conditions on children and their families. Yet measurement of participation has only been attempted relatively recently, with limited validation of existing measures available (Field, Miller, Ryan, Jarus, Abundo, 2016). One measure of children's participation, the Participation and Environment Measure for Children and Youth (PEM-CY: Coster et al., 2013a) includes concurrent measurement of environments and participation in light of the recognised relationship between these constructs (Anaby et al., 2014).

The PEM-CY is a 70-item, parent-report measure examining participation and environment in 3 settings: home, school and the community for children aged 5 to 17 years. Parents record children's participation in activities in each setting by indicating frequency of participation (Frequency), extent of participation (Involvement) and desire for change in either the frequency or extent of participation (Desire for Change). Features of children's environments are rated in relation to their Helpfulness or Availability. As such the PEM-CY is inclusive of multiple dimensions of participation and provides context-specific scoring. However, it is currently unclear if this inclusive and granulated approach to measurement of participation and environments produces valid information for clinical or research purposes.

The PEM-CY is recommended for clinical and research purposes, and has already been used widely including with children with and without disabilities (Bedell et al., 2013, Law et al., 2013, Coster et al., 2013b; Anaby, Law, Majnemer, & Feldman, 2016), after critical illness (Choong et al., 2015), for collaborative care planning (Khetani et al., 2015) and in health impact assessments (Khetani et al., 2014). As such, interest in the PEM-CY indicates the need for validation with construct validity being an important starting point.

One standard approach to determining construct validity is through determining convergent validity (what an instrument does measure), and divergent validity (what an instrument does not measure). Currently no evaluation of the convergent and divergent validity of the PEM-CY has been reported.

To date, psychometric evaluation of the PEM-CY is limited. A convenience sample of caregivers of children with and without disability (*n*= 578: Coster et al., 2011) reported low internal consistency across all settings and dimensions of Participation (Cronbach α >0.59) which the authors interpreted as acceptable for population studies, although not for individual assessments. Test-retest reliability ranged from moderate to good (ICC > 0.58). Known group validity (one of several methods of construct validity testing) has been examined in relation to presence of disability and has been found to consistently distinguish between disabled and non-disabled children (Coster et al., 2011; Jeong, Law, Stratford, et al., 2017; Law et al., 2013). However further examination of construct validity is recommended (Coster et al., 2011). To this end, the specific research aim of this study was to examine the construct validity of the PEM-CY through analysis of its convergent and divergent validity.

METHODS

This study is the quantitative component of a larger mixed methods study examining the construct validity of the PEM-CY (REMOVED FOR BLIND PEER REVIEW). In this component of the study construct validity was examined through analysis of the convergent and divergent validity of the PEM-CY with the Child and Adolescent Scale of Participation (CASP; Bedell, 2011b), Child and Adolescent Scale of the Environment (CASE; Bedell, 2011a) and Parenting Sense of Competence Scale (PSOC; Johnston and Mash, 1989) (see

figure 1). Convergent and divergent validity were tested using a priori predictions of correlations between measures.

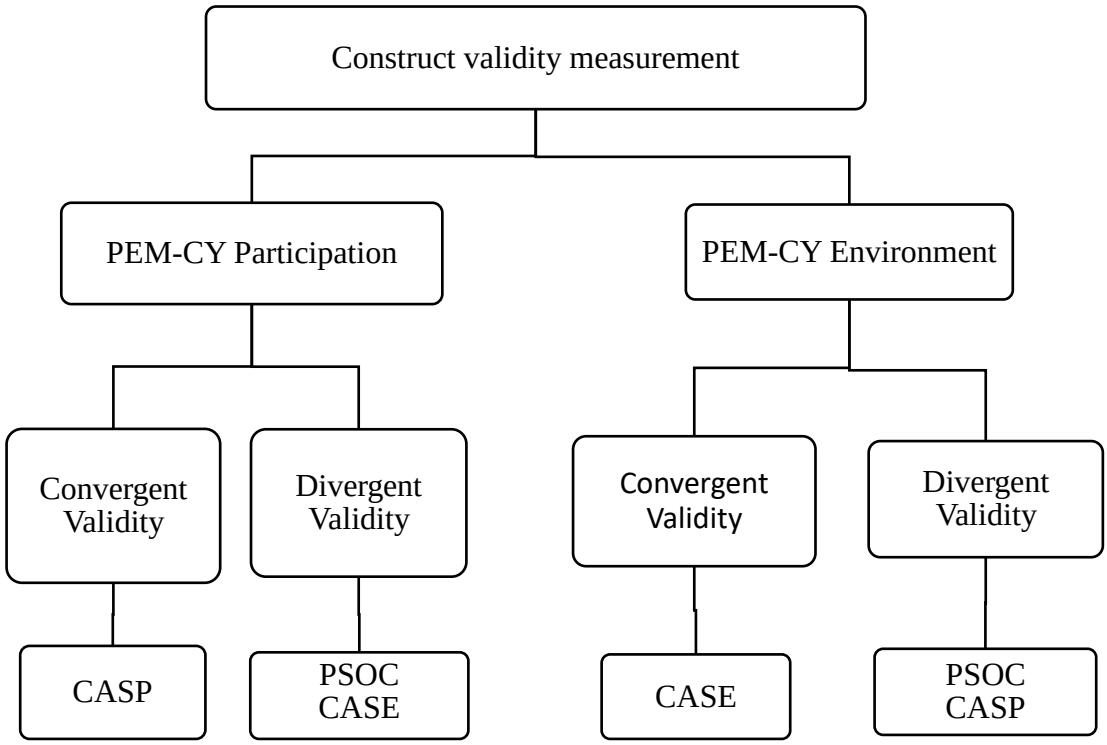

Figure 1. Measures compared for convergent and divergent validity analysis.

Participants were recruited from a convenience sample of university staff. It was not possible to determine the sampling frame because email invitations to the study were disseminated anonymously by a third party within the university. The sample size was calculated based on the power to detect a statistically significant correlation coefficient of at least 0.4 using Fisher's z-transformation. A minimum of 30 subjects was required to observe this level of correlation with type 1 error of 0.05. The largest sample size of 30 for the PSOC-PEMCY correlation was accepted as the minimum sample size required to identify the predicted correlations. Consistent with previous studies reporting construct validity of the PEM-CY (Coster et al., 2011) the inclusion criteria were: parents or guardians who had a

child aged between 5 and 17 years and were able to read English. There were no exclusion criteria. Participants were eligible to go in a draw for two NZ$100 grocery vouchers. Ethical approval for this study was obtained from the relevant university ethics committee (approval number 13/186).

Within the PEM-CY Participation is measured as (1) the average frequency of participation, (2) percentage of activities participated in, (3) average involvement and (4) percentage of activities in which change is desired. Environments are measured through (1) a count of supports and (2) barriers present; (3) ratings of helpfulness and resources; and (4) overall environmental support across settings. Scales used to measure each domain of participation and environment range from three to eight-point Likert scales. Author guidelines suggest the reporting of 27 summary scores (see Table 1; R. Teplicky, personal communication, August 20 2013 [PERMISSION GIVEN). No single one of these scores can be directly compared to an existing measure of participation. However, when Involvement and Frequency PEM-CY scores for each setting are combined they are arguably conceptually similar to the 'level' of participation described in the CASP. Specifically, in this study Participation composite scores were formed by multiplying each 'Frequency' score by the 'Involvement' score for each item (i.e., child activity) and dividing this by the total number of items rated within each setting to obtain a 'mean Participation score' for each setting. Composite scores were calculated for Overall Participation, Environmental Helpfulness and Environmental Resources. Obtaining a single indicator of participation for each setting reduced the number of scores from 27 down to 12 and created scores that were more directly comparable to corresponding instruments. All modifications to developer-recommended scoring are presented in Table 1. The Desire-for-Change summary was subsequently excluded from the analysis of construct validity since it measures 'satisfaction' rather than actual participation and no direct comparison measure was available.

Table 1. PEM-CY Recommended Scoring and Modified Summary Scores Derived from PEM-CY Raw Scores

PEM-CY Domain		Recommended PEM-CY Scores	Recommended PEM-CY Calculation	Modified PEM-CY calculation	Modified PEM-CY summary scores
Participation	Home	Frequency	Sum of all frequency responses/ number of home activities engaged in	Frequency x Involvement/ Number of home activities engaged in	Home participation
		Involvement	Sum of all involvement responses/ number of home activities engaged in		
		% activities	Number of home activities participated in/10 items x 100	NA	NA
		Desire for change	Sum of all desire for change responses/number of home activities engaged in x100	NA	NA
	School	Frequency	Sum of all frequency responses / number of school activities engaged in	Frequency x Involvement/ Number of school activities engaged in	School participation
		Involvement	Sum of all involvement responses / number of school activities engaged in		
		% activities	Number of school activities participated in/10 items x 100	NA	NA
		Desire for change	Sum of all desire for change responses /number of school activities done x100	NA	NA
	Community	Frequency	Sum of all frequency responses / number of community activities engaged in	Frequency x Involvement/ Number of community activities engaged in	Community participation
		Involvement	Sum of all involvement responses / number of community activities engaged in		

PEM-CY Domain		Recommended PEM-CY Scores	Recommended PEM-CY Calculation	Modified PEM-CY calculation	Modified PEM-CY summary scores
		% activities	Number of community activities participated in/10 items x100	NA	NA
		Desire for change	Sum of all desire for change responses / number of community activities engaged in x100	NA	NA
Environment	Overall	NA	NA	Home + School + Community Participation/3	Overall participation
	Helpfulness	Home	Sum of all home items/ maximum possible score x 100	NA	Home Helpfulness
		School	Sum of all school items/ maximum possible score x 100	NA	School Helpfulness
		Community	Sum of all community items/ maximum possible score x 100	NA	Community Helpfulness
		Overall	NA	Home+ school+ community helpfulness/3	Overall Environmental Helpfulness
	Resources	Home	Sum all home items/ maximum possible score x 100	NA	Home Resources
		School	Sum all school items/ maximum possible score x 100	NA	School Resources
		Community	Sum all community items/ maximum possible score x 100	NA	Community Resources
		Overall	NA	Home+ school+ community resources/3	Overall Environmental Resources
	Su nn	Home	Count of home environment items rated ‘usually helps’ or ‘usually yes’	NA	NA

PEM-CY Domain	Recommended PEM-CY Scores	Recommended PEM-CY Calculation	Modified PEM-CY calculation	Modified PEM-CY summary scores
	School	Count of school environment items rated ‘usually helps’ or ‘usually yes’	NA	NA
	Community	Count of community environment items rated ‘usually helps’ or ‘usually yes’	NA	NA
Barriers	Home	Count of home environment items rated ‘usually makes harder’ or ‘usually no’	NA	NA
Barriers	School	Count of school environment items rated ‘usually makes harder’ or ‘usually no’	NA	NA
Barriers	Community	Count of community environment items rated ‘usually makes harder’ or ‘usually no’	NA	NA
Overall support	Home	Sum of all home helpfulness and resources scores/ number of home items rated x100	NA	NA
Overall support	School	Sum of all school helpfulness and resources scores/ number of school items rated x100	NA	NA
Overall support	Community	Sum of all community helpfulness and resources scores/ number of community items rated x100	NA	NA
Total Number of Scores		27	NA	12

To test convergent validity of the PEM-CY Participation summary scores and the divergent validity of the PEM-CY Environment Summary scores, the CASP was chosen. Although no 'gold standard' participation or environment measures exist, the CASP is conceptually similar to PEM-CY-participation (Chien, Rodger, Copley, & Skorka, 2014), and has similar context-specific delineations. The CASP was also selected to test the divergent validity of the PEM-CY Environment sections because Environment scores were not expected to correlate strongly with participation, despite some degree of relationship being expected (Imms et al., 2016). The CASP is a 20-item parent-report measure of children's participation (ages 5-17 years). Subsections include: Home Participation (6-items), Neighbourhood and Community Participation (4 items), School Participation (5-items) and Home and Community Living Activities (5-items). Each item is rated using a 4-point Likert scale ranging from 'age expected' to 'unable' with an option of 'not applicable' for each item. Summary scores are calculated for each subsection and a total summary score. All CASP summary scores were calculated according to author guidelines.

The CASE was used to evaluate the convergent validity of PEM-CY Environment Summary scores and the divergent validity of the PEM-CY Participation Summary scores. The CASE contains 18-items rated on a 3-point Likert scale. All scoring of the CASE was in accordance with author guidelines.

Divergent validity of the PEM-CY was also examined using the PSOC as parental perceived competence was not expected to relate closely to children's participation. The PSOC is a 17-item parent self-report measure of perceived competence in parenting roles with well-established psychometric properties. Using the CASP and CAPE as both divergent and convergent measures reduced the assessment burden on participants.

To convert all measures and demographic questions into a single electronic questionnaire, participants answered items using an online secure commercial survey website (Survey Monkey™) with questions presented as similarly to measures' original paper formats. The order of measures were: PEM-CY, CASE, CASP and PSOC followed by demographic questions (PEM-CY comprises 122 out of the combined total of 190 questions presented to participants). Prior to recruitment (September, 2013), the electronic questionnaire was tested on parents of children aged 5-17 years (n=4) to ensure accurate interpretation of items and acceptability of online presentation. No changes to terminology or layout were required. Participants were advised the questionnaire would take 40 minutes to complete.

Statistical Analysis

The content of each PEM-CY summary score and each CASP, CASE or PSOC summary score were considered and expected correlations were predicted based on guidelines proposed by Burnand (1990) and on theoretical relationships. Scores from directly equivalent constructs and settings were predicted to have substantial levels of correlation (r_s >0.45; Burnand, 1990. Equivalent constructs with non-equivalent settings from each measure were expected to have a moderate correlation ($r_s \geq 0.3$). Non-equivalent constructs were predicted to have a low level of correlation ($r_s < 0.3$) with alternative summary scores. All predicted (hypothesised relationships that were tested are presented in Table 2). Because multiple scores in the comparison measures related to the construct being examined in the PEM-CY several correlations (range: 1 to 7) were examined for each PEM-CY summary score. Data were included if at least 50% of items were completed within one subsection of a single summary score. Construct validity (convergent and divergent) was analysed using Spearman's rank correlation due to non-normal distribution of data. For correlations to

illustrate acceptable construct validity, 75% of the predicted comparisons (see Table 2) needed to be met (Terwee et al., 2007).

Table 2 Predicted strength of relationship for convergent and divergent validity testing.

PEM-CY score		CASP home	CASP community	CASP school	CASP home and community living	CASP total	CASE	PSOC
Participation	Home	+++	++	++	++	++	+	+
	School	++	++	+++	++	++	+	+
	Community	++	+++	++	++	++	+	+
	Overall	++	++	++	++	+++	+	+
Environmental Helpfulness	Home	+	+	+	+	+	+++	+
	School	+	+	+	+	+	+++	+
	Community	+	+	+	+	+	+++	+
	Overall	+	+	+	+	+	+++	+
Environmental Resources	Home	+	+	+	+	+	+++	+
	School	+	+	+	+	+	+++	+
	Community	+	+	+	+	+	+++	+
	Overall	+	+	+	+	+	+++	+

+ weak (r<0.3); ++ moderate (r= 0.3 to 0.44); +++ substantial (r= 0.45 to 0.59); ++++ high (r>0.59) (Burnand, 1990)

Internal consistency of the original and modified PEM-CY scoring was examined in order to assess the internal consistency under both scoring conditions, given that assumptions of internal consistency underpin a meaningful interpretation of findings. Internal consistency of each measure was calculated using Cronbach's Alpha and were interpreted as acceptable if between α 0.70 and 0.95 (Portney and Watkins, 2000). Data were analysed using IBM SPSS, Version 22.0 (see Table 2).

Table 3 Characteristics of Parents (*n* = 76) and Children (*n*= 75)

Demographic	Parents (*n*)	Children (*n*)
Gender		
Male	10	39
Female	66	36
Age (years)		
Child		
5-9	-	26
10-14	-	30
15-17	-	19
Parent		
20-29	1	-
30-39	14	-
40-49	47	-
50-59	11	-
60-69	1	-
Ethnicity[1]		
NZ European	72	70
NZ Maori	2	7
Other	7	8
Level of Education		
High School	3	-
Tertiary	73	-
Travel time to nearest hospital		
0 to 30 minutes (urban)	65	-
31 to 60 minutes (rural)	11	-
Disability		
No disability	-	71
Disability	-	4

Note. [1]Participants could choose more than one ethnicity. NZ = New Zealand.

RESULTS

Two thirds (63%, 76/118) of parents who began the online questionnaire completed all measures during September, October, 2013. Missing data for some items resulted in variable sample sizes (N=67 to 100). Of the participating parents, 78% (76/97) provided demographic information about themselves and 77% (75/97) provided demographic information about their children (see Table 2 for detailed demographic information).

Most responders were mothers 87% (66/76) had tertiary qualifications 96% (73/76) and lived in an urban setting 86% (65/76). Children being reported on were almost evenly boys and girls, with most aged under 15 years 75% (56/75). Only a small proportion reported children as having a disability 5% (4/75).

Preliminary analysis of internal consistency revealed satisfactory internal consistency scores in less than half (8/15) of the PEM-CY summary scores using the original scoring procedure. Environment categories met the criteria for internal consistency more frequently (4/6) than Participation categories (3/9) (see Table 3). For the modified scores, internal consistency criteria were met for 3/6 categories, including all Overall scores (combined context scores).

Table 4 Internal consistency estimates of reliability (Cronbach's alpha) for PEM-CY

PEM-CY score		Author recommended scoring method α*(n)*			Modified scoring method used in this study α*(n)*
	Context	Frequency	Involvement	Desire for change	Frequency x Involvement
Participation	Home	0.20 (100)	0.58 (97)	0.67 (97)	0.51 (97)
	School	0.47 (90)	0.65 (87)	0.73* (86)	0.63 (87)
	Community	0.59 (89)	0.72* (84)	0.78* (87)	0.64 (84)
	Overall	NA[1]	NA	NA	0.79* (82)
Environmental helpfulness	Home		0.71*(95)		NA[2]
	School		0.63 (89)		NA
	Community		0.65 (85)		NA
	Overall		NA		0.77* (89)
Environmental resources	Home		0.77* (95)		NA
	School		0.82* (89)		NA
	Community		0.75* (84)		NA
	Overall		NA		0.93* (83)

Note.
[1] No overall scores reported by PEM-CY developers
[2] Scoring of environments scales was the same as recommended
* Internal consistency considered acceptable (α between 0.70 and 0.95)

Adequate convergent validity was observed only in the Environment Resources scores in relation to CASE summary scores. The PEM-CY Participation Summary scores showed very little evidence of convergent validity in relation to CASP scores (see Table 4).

Divergent validity was more often observed than convergent validity, especially for Environment Resources and Helpfulness summary scores. Most of the predicted relationships between Participation scores and CASE or PSOC were observed. Overall, >75% of predicted relationships were observed for some of the Environment Summary scores but for none of the Participation summary scores.

Table 5 Results of convergent and divergent validity testing for Modified PEM-CY scores.

PEM-CY score		CASP home N=77	CASP community N=77	CASP school N=70	CASP home and community living N=74	CASP total N=67	CASE N=75	PSOC N=70	Number of predicted correlations that were actually observed (%)[1]
Participation	Home	-0.29*	-0.20	**-0.32 ** **	-0.50 **	**-0.43****	**-0.12**	-0.30*	3 (43%)
	School	-0.14	-0.29*	-0.33**	**-0.34****	**-0.38****	**-0.27***	**-0.24***	4 (57%)
	Community	-0.25*	-0.39**	**-0.37****	**-0.39****	**-0.42****	**-0.20**	-0.30*	4 (57%)
	Overall	**-0.30****	**-0.40****	-0.45**	-0.51**	**-0.52****	**-0.27***	**-0.24***	5 (71%)
Environmental Helpfulness	Home	-0.73**	**-0.05**	**-0.07**	**-0.12**	**-0.13**	-0.28*	-0.36**	4 (57%)
	School	**-0.21**	**-0.01**	**-0.14**	**-0.06**	**0.01**	-0.30**	**0.00**	**6 (86%)**
	Community	**-0.08**	**-0.22**	**0.01**	**0.04**	**0.05**	-0.42**	**-0.16**	**6 (86%)**
	Overall	**-0.22**	**-0.03**	**-0.18**	**-0.08**	**-0.09**	-0.38**	**-0.12**	**6 (86%)**
Environmental Resources	Home	**-0.14**	**-0.22***	**-0.19**	**-0.18**	**-0.28***	**-0.53****	-0.41**	**6 (86%)**
	School	**-0.29***	**-0.27***	**-0.23***	**-0.26***	-0.34**	**-0.61****	-0.35**	5 (71%)
	Community	**-0.29***	**-0.25**	**-0.18**	-0.31**	-0.30*	**-0.66****	-0.35**	4 (57%)
	Overall	**-0.25***	**-0.26***	**-0.19**	**-0.26***	**-0.32****	**-0.67****	-0.40**	**6 (86%)**

Note. Bold type indicates correlation was as predicted. *p<0.05, **p<0.01

[1] Terwee et al (2007) guideline recommends that >75% of results confirm specific hypotheses for construct validity to be adequate

DISCUSSION

This study examined the construct validity of the PEM-CY through analysis of convergent and divergent validity with comparable measures of participation and environments and a peripherally related measure of parental sense of competence. Construct validity criteria were not met for the majority of PEM-CY summary scores, particularly in relation to Participation using the modified scoring. Explanations for these findings are discussed.

The internal consistency estimates of reliability (Cronbach's alpha) were not high which may have contributed to poor construct validity tests arising from excessive measurement error. As expected, alpha estimates were generally higher for the modified 'overall' summary scores due to the greater number of items contributing to those scores and are considered adequate (alpha >0.70) (Krabbe, 2017). Consistent with PEM-CY psychometrics reported by Coster et al, we observed lower alpha values for Frequency of Participation compared to Participation Involvement. It is possible that the frequency of engaging in activities is not sufficiently reliable, and may not justify our scoring approach (multiplying the frequency of an activity with the child's involvement in that activity to produce a single score). Alpha values are influenced by sample-homogeneity. The relatively homogenous sample in the present study (female, educated and urban) may have led to lower alpha values simply through less variation. The Environment Scores demonstrated generally much better internal consistency of reliability estimates, which was also observed by Coster et al. However, internal consistency may not be an appropriate indicator of the reliability of a measure of participation, given that items in such measures may not logically correlate highly because the items collectively *cause or form* the phenomenon (i.e., formative-type concepts) rather than being *reflective of it*.

The use of internal consistency testing for formative-type concepts can lead to instruments being inappropriately dismissed or criticised (Juniper et al., 1999). Measurement error in formative-type measures, such as measures of participation, is possibly better evaluated using test-retest procedures, rather than relying on the assumption of correlation between items.

Adequate construct validity, evidenced through predicted relationships between measures of similar and different concepts were not consistently observed in the PEM-CY Participation summary scores. Three possible explanations for this are (1) the predictions were incorrect; (2) the comparison measures do not quantify the intended constructs; and/or (3) the PEM-CY has insufficient construct validity.

First, predictions were based on the assumptions that the modified summary scores of the two measures within the same settings would be substantial; correlations of summary scores within different settings would be moderate; all PEM-CY Environment Summary scores would correlate substantially with the CASE total summary score and; all correlations with divergent measures would be insignificant. In light of findings, the relationship between participation and environmental factors appears more complex than was presumed and measures of more distinctive constructs are recommended for future divergent validity comparisons. Our apriori conceptually-based logic that multiplying PEM-CY Participation Frequency by Involvement in the attempt to obtain a comparable Participation score to the CASP (where Participation items are worded as 'level' of Participation) may have been incorrect. Nevertheless, this does not entirely explain why concurrent measures of participation failed to show expected correlation.

Second, the comparison measures may not have quantified participation and environments as they are proposed to. Variable results from validity studies of the CASP

(Bedell and Dumas, 2004, Bedell, 2009, Hwang et al., 2013) raise questions about its unidimensionality. This may relate to issues inherent to the construct of participation which will also affect the PEM-CY. Conversely, the relationship between the PEM-CY and the divergent validity measure (parental self-esteem; PSOC) may have been more strongly related to participation than was anticipated. However it remains a reasonable assumption that correlations between PSOC and PEM-CY scores would be lower than between PEM-CY scores and other measures of participation and environment (McDowell, 2006), which they were not.

Third, if findings indicate that the PEM-CY has insufficient construct validity this seems most strongly related to the challenges of measuring participation rather than errors in its development or scoring. Processes of development of the PEM-CY were clearly described and follow good practice based on classical test theory. Modified summary scores had consistently higher Cronbach's Alpha than the recommended scoring hence the deviation from the recommended scoring does not explain the finding of inadequate construct validity.

Consideration of whether items from PEM-CY (and other participation measures) are 'effect' or 'causal' indicators of participation may explain these findings (Fayers & Hand, 1997). We question the appropriateness of using a simple summated item score and the appropriateness of evaluation using Cronbach's alpha. The choice of effect or causal models of measurement design depend on the nature of the construct being measured. Where items reflect (i.e., are affected by) the underlying trait, it is expected that all items will vary with the amount of the trait of interest present. For example, intelligence could be measured by several different tests, the performance on which is supposed to vary by the level of intelligence of the test-taker. Other examples of effect constructs include upper limb strength and personality traits. As with the Rasch model, responses to items depend only on the difficulty

of the item and the amount of the trait possessed by the test-taker (Haley, Ludlow and Coster, 1993).

However, measures that are comprised of causal items, that is items that cause (i.e., influence) the construct are not expected to correlate highly, given they act independently on the construct (De Vet, Terwee, Mokkink et.al.,, 2011). For example, having sufficient income will influence quality of life, but having a good quality of life will not affect income, in the way that having higher intelligence will affect performance on a problem-solving task. Happiness and well-being are also examples of causal constructs. Likewise, PEM-CY items within Frequency, Involvement and Desire for Change domains may each *cause* Home Participation, but may not correlate with each other. It is very plausible that high frequency of participation in one particular activity will limit the time available to participate in another kind of activity, so the assumption of co-variation of items is logically difficult to sustain. Applied to measurement of participation and environments, the causal model may be more appropriate to measurement design and evaluation. Such instruments are more critically dependent on the appropriateness and relevance of their content, and are not appropriately evaluated by all aspects of classical test theory or item-response theory. In the case of a participation measure, it may be possible to create scenarios composed of a limited number of activities each with different levels of frequency and involvement. Using conjoint analysis, it is possible for children/parents to determine the relative level of participation between such scenarios and for the contribution of each activity frequency/involvement to be decomposed, thus creating a scoring system without making assumptions about how items should be summated.

If items from a ‘causal’ or ‘formative’ model (Krabbe, 2017) are summed together, the meaning of the summed score is not immediately clear. Often it is more clinically relevant to express the results of such instruments by the individual items rather than a summed score.

For example, the Glasgow Coma Scale (Teasdale and Jennett, 1974) is often expressed in the clinical setting by its individual components.

One useful approach that solves the issue of meaningful summation of items for formative/causal measures and which has been applied to quality of life measurement is the framework of conjoint analysis of discrete choice data (Arons et al., 2016). In such an approach, the key attributes that make up 'participation' are identified, together with different levels of those attributes, using, for example the coding structure of the International Classification of Functioning, Disability and Health (World Health Organisation, 2001). By asking relevant people to make repeated comparative judgements (choices) between two scenarios described only in terms of these attributes, as to which scenario demonstrates the 'most participation', it is possible to derive relative weights for each attribute-level that can then be applied as a scoring system for the instrument, even if culture-specific versions are needed. Such an approach has recently been used to develop a quality of life instrument for adults with dementia (Arons et al., 2016) and may advance valid measurement of participation. Collaborations between researchers from diverse disciplines (such as health, social and education sciences) with diverse methodological traditions may be beneficial in accelerating our understanding of the important, yet complex, constructs of participation and environment.

Limitations

Participants were predominantly white, female and educated. A stratified random sample from the general population would have provided more representative data however the sample reported here has similar characteristics to that of Coster et al. with the exception of child disability, thereby facilitating some comparison of findings. The alternative formatting of questions required for online distribution used in this study may have affected participants' responses. Future studies should use the electronic format now available

(https://www.canchild.ca/en/shop/2-pem-cy-participation-and-environment-measure-children-and-youth) given its arrangement is consistent with the paper version. The validity analyses are based on a modified scoring of the PEM-CY in order to allow comparison with another measure of participation. While analyses of reliability were slightly better for the modified PEM-CY scoring compared to the original scoring, it should be acknowledged that the findings for the Participation (but not Environment) scores are not based on the original PEM-CY scoring.

Findings from this study relate to convergent and divergent validity, which are only two forms of validity. Face validity is also important and findings in a parallel study suggest that clinicians and parents perceptions of the PEM-CY support its use in providing a clinical description of children's Participation and Environments, albeit with some caveats ([ref removed for blind peer review]). In addition, the individual items of the PEM-CY may be a useful clinical tool to describe the profile of Participation and Environmental issues for children, even if the summary scores may not be meaningful numerical quantities. Findings from this current study are therefore particularly cautionary regarding the use of the PEM-CY in research as an outcome measure, with acknowledgement that it may remain a valid tool for clinical use.

Conclusion:

Collectively, these findings may indicate that the modified summated scores derived from the PEM-CY do not measure children's participation in a valid way. As convergent validity testing pf the PEM-CY was not possible without some modification to scoring, the comparison of the PEM-CY and CASP was the best option available. We recommend that individual items of the PEM-CY be expressed for clinical use, and further efforts towards measurement of participation using a formative, rather than effect model, with use of conjoint analysis of discrete choice data being proposed as a way forward.

Key Points for Occupational Therapy:

- Participation is a multidimensional construct that requires alternative ways of designing instruments.
- When the PEM-CY scoring is modified to align with another measure of participation, it does not meet the criteria for construct validity.
- Conjoint analysis of discrete choice data, constructed from the key dimensions of participation, may provide a more useful approach to developing a measurement instrument for this formative-type concept.

REFERENCES

Anaby, D., Law, M., Coster, W., Bedell, G., Khetani, M., Avery, L., & Teplicky, R. (2014). The mediating role of the environment in explaining participation of children and youth with and without disabilities across home, school, and community. Archives of physical medicine and rehabilitation, 95(5), 908-917.

Anaby, D. R., Law, M. C., Majnemer, A., & Feldman, D. (2016). Opening doors to participation of youth with physical disabilities: An intervention study. *Canadian Journal of Occupational Therapy, 83*(2), 83-90. doi:10.1177/0008417415608653

Arons, A. M., Scholzel-Dorenbos, C. J., Olde Rikkert, M. G., & Krabbe, P. F. (2016). A simple and practical index to measure dementia-related quality of life. Value in Health, 19(1), 60-65. doi:doi:10.1016/j.jval.2015.07.011

Bedell, G. (2009). Further validation of the Child and Adolescent Scale of Participation (CASP). Developmental neurorehabilitation, 12(5), 342-351. doi:10.3109/17518420903087277

Bedell, G. (2011a). The Child and Adolescent Scale of Environment (CASE): Administration and scoring guidelines. Medford, MA: Author.

Bedell, G. (2011b). The Child and Adolescent Scale of Participation (CASP)–Administration and Scoring Guidelines. Unpublished manual. Available at: http://sites. tufts. edu/garybedell/measurement-tools.

Bedell, G., Coster, W., Law, M., Liljenquist, K., Kao, Y.-C., Teplicky, R., . . . Khetani, M. (2013). Community participation, supports, and barriers of school-age children with and without disabilities. Archives of physical medicine and rehabilitation, 94(2), 315-323.

Bedell, G., & Dumas, H. (2004). Social participation of children and youth with acquired brain injuries discharged from inpatient rehabilitation: A follow-up study. Brain Injury, 18(1), 65-82.

Burnand, B., Kernan, W. N., & Feinstein, A. R. (1990). Indexes and boundaries for "quantitative significance" in statistical decisions. Journal of Clinical Epidemiology, 43(12), 1273-1284. doi:http://dx.doi.org/10.1016/0895-4356(90)90093-5

Chien, C. W., Rodger, S., Copley, J., & Skorka, K. (2014). Comparative content review of children's participation measures using the international classification of functioning, disability and health-children and youth. *Archives of physical medicine and rehabilitation, 95*(1), 141-152. doi:10.1016/j.apmr.2013.06.027

Choong, K., Al-Harbi, S., Siu, K., Wong, K., Cheng, J., Baird, B., . . . Thabane, L. (2015). Functional recovery following critical illness in children: the "wee-cover" pilot study. Pediatric critical care medicine: a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies, 16(4), 310.

Coster, W., Bedell, G., Law, M., Khetani, M., Teplicky, R., Liljenquist, K., . . . Kao, Y. (2011). Psychometric evaluation of the participation and environment measure for children and youth. Developmental Medicine & Child Neurology, 53(11), 1030-1037.

Coster, W., Law, M., & Bedell, G. (2013). Participation and Environment Measure for Children and Youth (PEM-CY). Retrieved from https://www.canchild.ca/en/shop/2-pem-cy-participation-and-environment-measure-children-and-youth

Coster, W., Law, M., Bedell, G., Liljenquist, K., Kao, Y. C., Khetani, M., & Teplicky, R. (2013). School participation, supports and barriers of students with and without disabilities. Child: Care, Health and Development, 39(4), 535-543.

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: a practical guide*: Cambridge University Press.

Fayers, P., & Hand, D. (1997). Factor analysis, causal indicators and quality of life. *Quality of life research, 6*(2).

Field, D. A., Miller, W. C., Ryan, S. E., Jarus, T. & Abundo, A. 2016. Measuring participation for children and youth with power mobility needs: A systematic review of potential health measurement tools. *Archives of Physical Medicine and Rehabilitation*, 97**,** 462-477.e40.

Haley, S., Ludlow, L. & Coster, W. 1993. Pediatric Evaluation of Disability Inventory: clinical interpretation of summary scores using Rasch rating scale methodology. Physical Medicine & Rehabilitation Clinics of North America, 4, 529-540.

Hwang, A.-W., Liou, T.-H., Bedell, G. M., Kang, L.-J., Chen, W.-C., Yen, C.-F., . . . System, C. T. F. o. D. E. (2013). Psychometric properties of the Child and Adolescent Scale of Participation–traditional Chinese version. International Journal of Rehabilitation Research, 36(3), 211-220.

Imms, C., Adair, B., Keen, D., Ullenhag, A., Rosenbaum, P., & Granlund, M. (2016). 'Participation': a systematic review of language, definitions, and constructs used in intervention research with children with disabilities. *Developmental Medicine & Child Neurology, 58*(1), 29-38.

Jeong, Y., Law, M., Stratford, P., DeMatteo, C., & Missiuna, C. (2017). Measuring Participation of Children and Environmental Factors at Home, School, and in Community: Construct Validation of the Korean PEM-CY. *Physical and Occupational Therapy in Pediatrics*, 1-13. doi:10.1080/01942638.2017.1280870

Johnston, C., & Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. Journal of Clinical Child Psychology, 18(2), 167-175.

Juniper, E., Guyatt, G., Ferrie, P., & King, D. (1999). Development and validation of a questionnaire to measure asthma control. European Respiratory Journal, 14(4), 902-907.

Khetani, M., Cliff, A., Schelly, C., Daunhauer, L., & Anaby, D. (2015). Decisional support algorithm for collaborative care planning using the Participation and Environment Measure for Children and Youth (PEM-CY): a mixed methods study. Physical & Occupational Therapy in Pediatrics, 35(3), 231-252.

Khetani, M., Marley, J., Baker, M., Albrecht, E., Bedell, G., Coster, W., . . . Law, M. (2014). Validity of the Participation and Environment Measure for Children and Youth (PEM-CY) for Health Impact Assessment (HIA) in sustainable development projects. Disability and health journal, 7(2), 226-235.

Krabbe, P. F. (2017). The measurement of health and health status: concepts , methods and applications from a multidisciplinary perspective. Amsterdam: Academic Press Elsevier.

Law, M., Anaby, D., Teplicky, R., Khetani, M., Coster, W., & Bedell, G. (2013). Participation in the home environment among children and youth with and without disabilities. The British Journal of Occupational Therapy, 76(2), 58-66.

McDowell, I. 2006. Measuring health: a guide to rating scales and questionnaires, Oxford university press.

Portney, L., & Watkins, M. (2000). Foundations of Clinical Research (2nd ed.). Sydney: Prentice-Hall.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. Lancet, 2(7872), 81-84.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., . . . de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology, 60(1), 34-42.

World Health Organisation. (2001). International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organisation.