J Rehabil Med 2011; 43: 631–637

ORIGINAL REPORT

CROSS-CULTURAL VALIDITY OF A NATURALISTIC OBSERVATIONAL ASSESSMENT OF CHILDREN'S HAND SKILLS: A STUDY USING RASCH ANALYSIS

Chi-Wen Chien, PhD[1], Ted Brown, PhD[2] and Rachael McDonald, PhD[2]

From the [1]Occupational Therapy Division, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland and [2]Department of Occupational Therapy, School of Primary Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University–Peninsula Campus, Frankston, Victoria, Australia

Objective: **To investigate cross-cultural validity of a newly developed Assessment of Children's Hand Skills (ACHS) in children with and without disabilities.**

Subjects: **One group of 138 Australian children and 134 Taiwanese children, age range 2−12 years (a total of 149 typically developing children and 123 children with disabilities).**

Methods: **Rasch model analysis was used to evaluate rating scale performance, person-fit, and item unidimensionality of pooled data from the ACHS. Rasch-based differential item functioning was used to evaluate differences in item difficulties between cultural groups.**

Results: **The appropriateness of the ACHS 6-level rating scale was confirmed in the pooled analysis. All 22 activity items and 19 of 20 hand skill items in the ACHS formed a unidimensional scale. The ACHS exhibited reasonable response patterns when applied to the composite sample of Australian and Taiwanese children. Differential item functioning was found in 7 activity items and 7 hand skill items.**

Conclusion: **The ACHS can be used to assess and compare children's hand skills across Australian and Taiwanese cultural contexts with confidence. Adjustment for differential item functioning may be used when pooling or comparing ACHS data from these cultural groups.**

Key words: motor skills; children; psychometrics; outcome assessment (healthcare).

J Rehabil Med 2011; 43: 631–637

Correspondence address: Chi-Wen Chien, Occupational Therapy Division, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland, 4072, Australia. E-mail: c.chien@uq.edu.au

Submitted December 12, 2010; accepted April 14, 2011

INTRODUCTION

Most activities in daily life typically involve hand use; however, mastery of hand skills is achieved gradually between childhood and adulthood. Difficulties in hand use in children could therefore lead to restricted participation in play, school, social, and self-care activities (1–3).The assessment of children's hand skills is an important skill for clinicians to develop and is vital in evaluating intervention programmes for children with physical or intellectual disabilities.

A number of instruments are available to assess children's hand skills for clinical and research purposes, such as the fine motor subscale/component of the Peabody Developmental Motor Scales (4) or the Movement Assessment Battery for Children (5). Most of these instruments focus on whether a child can (or cannot) complete a set of standardized tasks (e.g. buttoning a 5-button strip) with his or her hands in standardized environments. This type of assessment gives an indication of the child's hand skill capacities that allow comparison of his or her performance under comparable circumstances. However, hand skill capacities do not necessarily generalize to the child's actual performance of activities of daily living (ADL) in natural contexts (e.g. buttoning own clothes, putting toothpaste on a brush, and combing hair) (6). The distinction between hand skill capacity (what a child *can do*) and real-life performance (what he/she *actually does do* in daily life) is also recognized with two different qualifiers in the Activity and Participation component of the International Classification of Functioning, Disability and Health (ICF) (7). Thus, assessment of hand skills performed in real-life contexts has been increasingly emphasized so as to promote children's participation in daily life (8, 9).

The newly developed Assessment of Children's Hand Skills (ACHS) (10, 11) is one of few instruments that capture children's real-life hand skill performance. The ACHS utilizes observation to assess the effectiveness of hand skills of 2 to 12-year-old children when engaged in play, school-related, or self-care activities within natural contexts. Observation of real-life contexts provides an ecological base for assessment that takes into consideration a child's unique characteristics (e.g. ethnicity, habits, or experiences) and environmental factors (e.g. physical or cultural) (7, 12). The potential of the ACHS for use in cross-cultural settings is therefore expected. Furthermore, the ACHS can be also used with most groups of preschool- and school-age children with and without disabilities (10, 11).

Content validity of the ACHS has been established through literature review and several rounds of international expert review (10, 11). The ACHS research version includes 22 typical activities that involve hand use, present minimal environmental requirements, and are cross-culturally compatible. It also includes 20 common hand skill items that children may exhibit in daily life, such as reaching for, grasping, and manipulating objects. The hand skill activities are used as assessment

contexts to observe and rate effectiveness of hand skills. The assessment construct (i.e. children's hand skills performed in natural contexts) has been subsequently validated using a Rasch measurement model, and its unidimensionality was deemed acceptable after removing one hand skill item (11, 13). Preliminary test-retest reliability estimates of the ACHS (Pearson's r coefficient = 0.85, $p < 0.01$) and inter-rater reliability ($r = 0.79$, $p < 0.01$) (10, 11) have been reported.

There is currently no validity evidence supporting whether the ACHS can be used in a consistent manner in cross-cultural contexts. Given that the ACHS provides ecologically-based measurement in cultural contexts, and that multinational research projects are increasing in number, it is important to examine whether ACHS data from different countries can be pooled to facilitate cross-national comparison of children's hand skill performance. Furthermore, previous studies have found that different cultural traditions, expectations, and environmental contexts have potential impact on children's motor/hand skills (14, 15), such as children's involvement in specific types of activities, their skilfulness for certain motor/hand skills, and varied rates of motor/hand skill development. Therefore, it is interesting to explore whether children from differing cultural backgrounds perform hand skills similarly in natural contexts. That is, whether these ACHS hand skill items and activity items present inherent advantages or disadvantages to certain cultural groups of children.

The present study examined the cultural invariance of the ACHS with a composite sample of Australian and Taiwanese children. Rasch-based differential item functioning (DIF) was examined to confirm whether the ACHS hand skill items and activity items were performed in a similar manner across these groups. The presence of DIF would indicate that certain test items do not reflect the same levels of the construct in different participant groups, which may limit data combination or comparison between groups (16).

METHODS

Participants

Two sets of data were collected as part of two previous studies examining the validity of the ACHS in two groups of children; one in Australia and one in Taiwan (11, 13). Each group included both typically developing children and those with known disabilities. Typically developing children were full-term, of mean birthweight, and free of impairments according to parent or caregiver report. Inclusion criteria for children with disabilities were that they had a formal medical or rehabilitation diagnosis of neuromuscular disorder (e.g. cerebral palsy and brachial plexus birth palsy), autism spectrum disorder, developmental disorder, or genetic/chromosome disorder (e.g. Down's syndrome) and had difficulty using their hands in play, educational, or activities of daily living, based on parent report.

Ethical approval for the study was granted by the Monash Standing Committee on Ethics in Research Involving Humans, the Victorian Department of Education and Early Childhood Development, and committees of the participating childcare centres, schools, and hospitals/clinics. Written consent was obtained from the participants' parents or caregivers.

Measures

The ACHS (10, 11) is a naturalistic observational tool that assesses the effectiveness of children's use of hand skills when engaged in meaningful activities in natural contexts. It is comprised of 20 hand skill items and 22 activity items. The hand skill items represent 6 types of skill categories: manual gesture (1 item), body contact hand skills (1 item), arm-hand use (7 items), adaptive skilled hand use (5 items), bimanual use (3 items), and general quality (3 items). The 22 activity items included in the ACHS range across 3 childhood domains of leisure/play (8 items), school/education (8 items), and activities of daily living (6 items). These activities cover a range of potential difficulty/ability levels and are used to elicit hand skill performance in children with a range of clinical conditions and at different age levels (e.g. 2–12 years). Brief descriptions of the activity and hand skill items of the ACHS are provided in Appendix SI (available from https://doi.org/10.2340/16501977-0827).

The ACHS requires raters to observe and score hand skill items in certain activities. Scoring is based on the effectiveness of the child's use of individual hand skills regardless of whether one or two hands are involved in activity performance. A 6-level Likert rating scale is used across all hand skill items with general meanings of 1 = very ineffective, 2 = ineffective, 3 = slightly ineffective, 4 = slightly effective, 5 = effective, and 6 = very effective. The lower 3 levels indicate inefficient hand skill performance that disrupts activity performance (e.g. unacceptable delay, excessive exertion or need for others' assistance), whereas the top 3 levels indicate that hand skills support activity completion. Detailed guidelines and examples for each rating scale across hand skill items can be found in the test manual (11). However, not all of the hand skill items have to be scored (10) and, if a child does not use certain hand skills in the activity, these items can be reported as "not observed". Given that that the ACHS is a new instrument, raters who use it may need training to administer and score the items reliably (11).

In addition, a parent-report questionnaire is included as part of the ACHS to promote client- or family-centred assessment and select appropriate activities and environment(s) for individual children's observations (10, 11). The questionnaire elicits parents' perceptions regarding which of the 22 ACHS activities present the right level of challenge for the child being observed. A 3-level Likert scale (i.e. extremely difficult, difficult, and not difficult) is used for the parents to rate their child's hand skill proficiency at carrying out these activities. Activities not attempted by their child in the last 3 months are marked as "not applicable". The information obtained assists in determining suitably challenging activities (i.e. those rated as "difficult") as well as the environments (e.g. home, school, or other) that provide the most appropriate opportunities for the ACHS observations to occur.

Procedure

Prior to the ACHS observations, the consenting parents or caregivers of eligible participants completed the aforementioned parent-report questionnaire. On the basis of the obtained information, one trained rater (i.e. the first author) observed and scored each participant's hand skill performance while undertaking at least two appropriate challenging activities. The selection of at least two activities was based on evidence from the well-established Assessment of Motor and Process Skills (12) that observation of more than one activity yields a more valid and reliable estimation of a child's performance. The rater previously demonstrated preliminary test-retest reliability (10, 11).

To observe hands skills in natural contexts, the potential date/time for each child to perform the chosen activities in his or her typical routine was obtained from the parents/caregivers/teachers. During the provided date/time, the rater came and sat in a place close to the child. Necessary moves/interruptions were minimized so as not to intrude while observing the child performing the activities spontaneously or motivated by the parents or teachers. Each activity was observed for a maximum time period of 10 min, and most observations were conducted within the children's real-life environments, such as homes, childcare centres, kindergartens, or school settings. Scoring was completed simultaneously as the child performed each activity or immediately after the activity was completed.

Data analysis

DIF analysis based on Rasch measurement model was used to examine cultural invariance properties of the ACHS. Within the Rasch model's

framework, items that measure a unidimensional construct in a test should maintain the same difficulty levels irrespective of individuals' characteristics including cultural difference (17). This expectation guides examination of DIF in test items across different cultural groups. To perform DIF analysis, the sample size for each group requires a minimum of 30 participants and, more reliably, approximately 150 in order to generate stable item difficulty calibration (18).

Rasch-based DIF analysis was carried out with the pooled sample of Australian and Taiwanese groups of children. Data were analysed with the Many-Faceted Rasch model (based on rating scale approach) using the FACETS computer software version 3.65.0 (19). Five facets were included: (*i*) children's abilities; (*ii*) the difficulty levels of hand skill items; (*iii*) the challenge levels of activity items; (*iv*) the levels of rating scale categories; and (*v*) the children's cultural backgrounds. The fifth (cultural) facet was set as a dummy facet, since this enabled the DIF analysis, but was not substantially included in the measurement estimation.

Fit with the Rasch model: rating scale, unidimensionality, and person response validity

We first investigated whether the pooled sample data fulfilled Rasch model expectations for optimized rating scale, unidimensionality, and person response validity, which must demonstrate acceptable fit before DIF can be investigated (17, 20). Therefore, the initial analysis focused on examining the appropriateness of the 6-level ACHS rating scale in hand skill items (19, 21).

Goodness-of-fit statistics were used to examine person- and item-fit to the Rasch model expectations. Infit and outfit mean square (MnSq) of < 1.4 in combination with standardized Z values (Zstd) of < 2.0 have been regarded as acceptable for a clinical observational assessment (8, 21) and were adopted in this study. An overall rate of 95% of children exhibiting goodness-of-fit indicates acceptable person response validity and, similarly, acceptable fit of 95% for the activity and hand skill items in the ACHS provides evidence of test unidimensionality (8, 12). In addition, a principal component analysis (PCA) of the standardized residuals was conducted to confirm whether the remaining ACHS hand skill items reflected a single construct of children's hand skills that are performed in natural contexts (22, 23). The criteria used were that the Rasch identified construct (principal component) explained > 60% of the variance and the residuals were randomly distributed and uncorrelated (i.e. eigenvalue size of less than 3 or less than 5% of the variance) (21, 22). The PCA of the item residuals was conducted using WINSTEPS 3.64.0 software (24).

Rasch analysis and adjustment for differential item functioning

The presence of cultural DIF was investigated by performing *t*-tests that are in-built as part of the FACETS (19). To maintain comparability of item difficulties across cultural groups, they were estimated by anchoring person-ability measures in each group to those of the pooled sample. DIF was identified by applying two recommended criteria simultaneously: (*i*) a DIF contrast such as the difference between separate item-difficulty measures of more than 0.5 logits; and (*ii*) statistically significant difference with *t* statistics (25, 26). Due to the number of repeated *t*-tests undertaken, Bonferroni corrections (27) were applied to adjust *p*-values to 0.0022 and 0.0025 for the ACHS activity and hand skill items, respectively.

Items displaying DIF were adjusted by using the approach recommended by Tennant et al. (17) to make group comparison possible. In this approach, the DIF items were split (or sub-divided) into two group-specific items with one item for Australian children and one item for Taiwanese children. An overall analysis was then performed with the items with no DIF, combined with the group-specific items for the items that showed DIF. Non-fitting items were removed and those exhibiting DIF were also adjusted until a fitting model was achieved and no DIF items were found (17).

Finally, we examined whether the adjustment for DIF had an impact on the children's overall ability estimates. The ability estimates were produced by Rasch analysis to represent each child's summation score in the ACHS, which was synthesized from the scoring of hand skill items when the child performed several chosen activities. The ability estimates are expressed by logit scores (which are interval-level measures) and indicate children's composite hand skill performance. Therefore, Rasch-derived person-ability estimates for the Australian and Taiwanese children were calculated by using the versions of the ACHS with and without adjusting DIF items. Minimal relevance was defined as the mean and individual differences between the unadjusted and adjusted person-ability estimates less than 0.5 logits and the Pearson's *r* correlation greater than 0.98 (28, 29).

RESULTS

A total of 272 children (138 from Australia and 134 from Taiwan) were included in the pooled sample for the study. The characteristics and ACHS results of each Australian and Taiwanese participant group are given in Table I. There were no significant differences between the two groups in age, gender, and handedness based on *t*-test or χ^2 comparison. However, the Taiwanese group included more children with disabilities than the Australian group and they demonstrated significantly lower ACHS logit scores than the Australian group (Table I).

A total of 1468 observational assessments were undertaken with 764 assessments from the Australian group and 704 observations from the Taiwanese group. The Australian and Taiwanese samples included assessments that were evenly distributed amongst the 3 domains of leisure/play (31.3% and 35.1%, respectively), school/education (40.6% and 37.5%), and activities of daily living (28.1% and 27.4%). Approximately 60.0% and 53.6% of the assessments were completed on typically developing children, and 53.6% and 62.5% were

Table I. *Characteristics and assessment results of the Australian and Taiwanese groups of children*

Variables	Australian sample (*n* = 138)	Taiwanese sample (*n* = 134)	Pooled sample (*n* = 272)
Gender, *n* (%)			
Boys	75 (54.3)	87 (64.9)	162 (59.6)
Girls	63 (45.7)	47 (35.1)	110 (40.4)
Age, months, mean (SD)	83.1 (34.4)	76.1 (31.1)	79.6 (32.9)
Age, years, *n* (%)			
2−4	46 (33.3)	45 (33.6)	91 (33.5)
5−7	41 (29.7)	53 (39.5)	94 (34.5)
8−10	33 (23.9)	25 (18.7)	58 (21.3)
11−12	18 (13.1)	11 (8.2)	29 (10.7)
Diagnosis, *n* (%)			
Non-disabled	85 (61.5)	64 (47.8)	149 (54.8)
Autism spectrum disorder	19 (13.8)	16 (11.9)	35 (12.9)
Developmental disorders	15 (10.9)	31 (23.1)	46 (16.9)
Genetic/chromosome disorders	12 (8.7)	6 (4.5)	18 (6.6)
Neuromuscular disorders	7 (5.1)	17 (12.7)	24 (8.8)
Handedness, *n* (%)			
Right	77 (55.8)	89 (66.4)	166 (61.0)
Left	16 (11.6)	14 (10.5)	30 (11.0)
Undetermined	44 (31.9)	31 (23.1)	75 (27.6)
Unreported	1 (0.7)	−	1 (0.4)
Assessment of Children's Hand Skills			
Mean (SD)	2.6 (3.7)	1.7 (3.0)	−
Range	–7.0−11.0	–4.7−10.2	−

−: not applicable; SD: standard deviation.

completed on boys. For each activity category, at least 20 observational assessments were included in the Australian group (mean = 34.7; standard deviation (SD) = 13.6) and the Taiwanese group (mean = 32.0; SD = 8.9).

Fit with the Rasch model: rating scale, unidimensionality, and person response validity

Rasch analysis revealed sufficient frequency counts (i.e. > 10 observations) and category fit (i.e. category outfit MnSq values < 2.0) in each of the 6 levels. None of the rating categories exhibited disordered step calibrations. This indicates that the ACHS 6-level rating scale was suitably used in the pooled sample, and hence no revision or modification was made.

With regards to person response validity, 31 out of 272 children (11.4%) exhibited misfit to the Rasch model expectations. This number exceeded 5%, which is Fisher's recommended criterion (12, 30). Therefore, 9 of the most extremely misfitting children were removed by using infit/outfit MnSq > 2.0 and Zstd > 2.0, which indicated their responses varied (or misfit) more than 100% from what the Rasch model predicted (21). Seven of the 9 extremely misfitting children had disabilities (e.g. autism spectrum disorder, developmental disorder, or cerebral palsy), and 2 were typically developing children. Considering the potential threat for the ACHS item-fit analysis, the 9 extremely misfitting children were removed from the analysis sample. The reduced sample of 263 children (8.1% misfitting) was used in the subsequent analyses.

The Rasch analysis results of the ACHS activity items and hand skill items in the pooled sample are shown in Tables II and III. All 22 activity items were found to demonstrate acceptable fit values, whereas only the Catching item out of 20 hand skill items exhibited misfit to Rasch model. Since the Catching hand skill item also did not have acceptable goodness-of-fit in the individual Australian and Taiwanese samples (11, 13), this item was deleted. Re-analysis confirmed that there were no additional misfitting activity items or hand skill items. PCA of the ACHS (after removal of the Catching item) revealed that 92.8% of total variance was explained by the Rasch-derived measures and only 0.8% and 2.2 eigenvalue was accounted for by the second major component. Therefore, the unidimensionality of 22 activity items and 19 hand skill items was deemed acceptable for the combined sample of Australian and Taiwanese children.

Rasch analysis and adjustment for differential item functioning

The results of DIF analysis revealed that 7 activity items and 7 hand skill items of the ACHS had significantly large difficulty calibration differences between Australian and Taiwanese children (see Tables IV and V). These 14 activity and hand skill items were subsequently split into two group-specific items and the analyses were repeated. The Using computer activity item specific to Taiwanese children and two hand skill items (Manual gesture and In-hand manipulating items) specific to Australian children were further identified as misfitting and were removed. The subsequent examination revealed that all of the remaining 15 common and 13 culture-specific activity items as well as 12

Table II. *Fit statistics for the 22 activity items of the Assessment of Children's Hand Skills*

Activity items	Measure	SE	Infit MnSq	Infit Zstd	Outfit MnSq	Outfit Zstd
Leisure and Play domain						
Construction (blocks)	–0.60	0.07	0.9	–2.6	0.9	–1.8
Puzzle	–0.10	0.06	0.9	–2.4	0.9	–2.0
Stringing beads	–0.02	0.07	0.9	–2.8	0.9	–2.4
Catching, throwing, & hitting/batting	0.30	0.05	1.2	5.1	1.2	3.5
Card game	0.62	0.08	1.0	–0.2	1.1	1.0
Playdough/clay	–0.28	0.07	0.8	–4.3	0.8	–4.3
Folding paper	0.95	0.07	1.0	0.1	1.0	–0.7
Handling money	0.33	0.07	0.9	–2.0	1.0	–0.8
School/Education domain						
Turning book	–1.19	0.07	1.0	0.8	1.1	2.0
Drawing and/or colouring	0.17	0.05	0.9	–1.8	0.9	–1.2
Writing & copying	–0.26	0.06	1.0	–0.9	1.0	–0.6
Cutting	0.57	0.06	1.0	–0.7	1.0	0.3
Pasting	0.70	0.07	0.9	–3.1	0.8	–2.7
Using computer	0.78	0.08	1.2	3.5	1.2	2.0
School tool use (ruler)	0.97	0.08	1.1	1.4	1.0	0.3
Putting on backpack	–0.73	0.09	1.2	2.5	1.1	1.4
Activities of Daily Living domain						
Drinking	–1.13	0.07	1.1	2.7	1.2	2.6
Eating	–0.50	0.05	1.0	–1.3	0.9	–1.0
Dressing upper body	0.55	0.07	1.1	1.2	1.0	0.7
Putting on socks and shoes	–0.29	0.07	1.0	0.3	1.0	–0.4
Washing hands	–1.09	0.07	1.2	3.8	1.2	1.5
Brushing teeth	0.24	0.08	1.0	0.1	0.9	–1.1

SE: standard error; MnSq: mean square; Zstd: standardized Z value.

Table III. *Fit statistics for the 20 hand skill items of the Assessment of Children's Hand Skills*

Hand skill items	Measure	SE	Infit MnSq	Infit Zstd	Outfit MnSq	Outfit Zstd
Manual gesture	–0.95	0.12	1.2	2.2	1.2	2.0
Body contact hand skills	–0.57	0.11	1.0	–0.5	1.0	–0.4
Reaching	–2.14	0.06	0.7	–9.0	0.6	–6.1
Turning	–0.13	0.05	0.9	–3.1	0.9	–2.6
Carrying	–1.04	0.06	0.8	–4.8	0.8	–4.0
Throwing	1.88	0.05	1.2	1.0	1.4	2.0
Catching	**4.58**	**0.18**	**2.2**	**6.5**	**2.4**	**6.8**
Moving	0.76	0.06	1.1	3.0	1.2	3.4
Stabilising	–0.41	0.06	1.3	6.3	1.2	4.5
Grasping	–1.24	0.06	0.9	–1.9	1.0	–0.9
Holding	–2.99	0.06	0.8	–4.1	0.8	–2.5
In-hand manipulating	2.07	0.07	1.2	3.3	1.2	3.3
Releasing	–1.58	0.06	0.9	–3.3	1.0	–0.8
Isolated finger movement	0.08	0.05	1.0	0.1	1.1	1.1
Transferring	–0.92	0.07	0.7	–5.9	0.8	–3.8
Using both hands simultaneously	–0.43	0.08	1.2	2.8	1.2	2.5
Using both hands cooperatively	1.25	0.05	1.0	0.4	1.0	0.6
Accuracy	0.01	0.05	1.2	4.8	1.3	5.6
Pace	–0.11	0.05	1.2	4.9	1.2	5.0
Movement quality	1.85	0.05	0.8	–4.5	0.9	–2.6

Misfitting item in bold type.
SE: standard error; MnSq: mean square; Zstd: standardized Z value.

Table IV. *Differential item functioning for the 22 activity items of the Assessment of Children's Hand Skills*

Activity items	Australian sample Measure (SE)	Taiwanese sample Measure (SE)	Difference	Calculated *t* value	*p*-value
Leisure and Play domain					
Construction (blocks)	–0.48 (0.10)	–0.71 (0.09)	0.22	1.63	0.1029
Puzzle	–0.25 (0.09)	0.02 (0.08)	–0.27	–2.21	0.0274
Stringing beads	–0.08 (0.11)	0.01 (0.08)	–0.09	–0.62	0.5340
Catching, throwing, & hitting/batting	**0 (0.08)**	**0.72 (0.08)**	**–0.73**	**–6.41**	**<0.0001**
Card game	0.82 (0.12)	0.48 (0.10)	0.34	2.10	0.0365
Playdough/clay	–0.23 (0.10)	–0.32 (0.10)	0.10	0.67	0.5020
Folding paper	**1.29 (0.11)**	**0.70 (0.10)**	**0.59**	**4.05**	**0.0001**
Handling money	0.55 (0.09)	0.11 (0.10)	0.44	3.26	0.0012
School/Education domain					
Turning book	–1.11 (0.10)	–1.29 (0.10)	0.18	1.27	0.2044
Drawing and/or colouring	0.06 (0.07)	0.31 (0.08)	–0.25	–2.51	0.0123
Writing & copying	–0.17 (0.10)	–0.31 (0.08)	0.14	1.14	0.2552
Cutting	**0.87 (0.07)**	**0.19 (0.09)**	**0.67**	**5.95**	**<0.0001**
Pasting	0.58 (0.09)	0.84 (0.09)	–0.26	–1.98	0.0477
Using computer	**0.50 (0.11)**	**1.12 (0.12)**	**–0.61**	**–3.69**	**0.0002**
School tool use (ruler)	0.93 (0.12)	1.01 (0.10)	–0.09	–0.56	0.5740
Putting on backpack	–0.56 (0.12)	–0.94 (0.12)	0.37	2.19	0.0291
Activities of Daily Living domain					
Drinking	**–1.56 (0.09)**	**–0.63 (0.10)**	**–0.92**	**–6.62**	**<0.0001**
Eating	–0.46 (0.07)	–0.54 (0.07)	0.08	0.81	0.4163
Dressing upper body	**0.25 (0.10)**	**0.84 (0.09)**	**–0.59**	**–4.30**	**<0.0001**
Putting on socks and shoes	**0.09 (0.11)**	**–0.50 (0.08)**	**0.58**	**4.08**	**<0.0001**
Washing hands	–0.97 (0.09)	–1.36 (0.13)	0.40	2.57	0.0105
Brushing teeth	0.34 (0.13)	0.18 (0.11)	0.16	0.95	0.3422

Bonferroni correction was applied giving a significant *p*-value of 0.0022 (0.05/22) for identification of statistical differences; activity items exhibiting culture-DIF are in bold type.
SE: standard error.

common and 12 culture-specific hand skill items were free of cultural DIF and met the Rasch model requirements.

The impact of the splitting adjustment for culture-DIF items on the children's ability estimates was investigated. The children's ability measures with and without adjusting for culture-DIF exhibited a small mean difference (e.g. 0.09 logits with a SD of 0.28 logits). Fewer than 5% of the children (i.e. 3 children with disabilities and 10 typically developing children)

Table V. *Differential item functioning for the 19 hand skill items of the Assessment of Children's Hand Skills*

Hand skill items	Australian sample Measure (SE)	Taiwanese sample Measure (SE)	Difference	Calculated *t* value	*p*-value
Manual gesture	**–1.11 (0.16)**	**–0.28 (0.17)**	**–0.83**	**–3.56**	**0.0004**
Body contact hand skills	**–0.64 (0.15)**	**0.15 (0.18)**	**–0.79**	**–3.40**	**0.0007**
Reaching	–1.96 (0.09)	–1.87 (0.08)	–0.09	–0.78	0.4362
Turning	**–0.19 (0.08)**	**0.38 (0.07)**	**–0.56**	**–5.32**	**<0.0001**
Carrying	–0.93 (0.08)	–0.69 (0.08)	–0.24	–2.13	0.0334
Throwing	2.11 (0.23)	2.19 (0.27)	–0.09	–0.25	0.8059
Catching[a]					
Moving	1.11 (0.08)	0.91 (0.08)	0.20	1.85	0.0647
Stabilising	**0.14 (0.08)**	**–0.47 (0.08)**	**0.61**	**5.21**	**<0.0001**
Grasping	–0.96 (0.08)	–1.05 (0.08)	0.09	0.80	0.4229
Holding	**–2.31 (0.09)**	**–3.19 (0.09)**	**0.88**	**6.93**	**<0.0001**
In-hand manipulating	**2.06 (0.10)**	**2.60 (0.10)**	**–0.54**	**–3.86**	**0.0001**
Releasing	–1.43 (0.08)	–1.27 (0.08)	–0.16	–1.38	0.1670
Isolated finger movement	0.12 (0.08)	0.52 (0.08)	–0.40	–3.69	0.0002
Transferring	**–0.40 (0.10)**	**–1.04 (0.10)**	**0.64**	**4.40**	**<0.0001**
Using both hands simultaneously	–0.03 (0.10)	–0.41 (0.12)	0.38	2.33	0.0199
Using both hands cooperatively	1.60 (0.08)	1.42 (0.08)	0.17	1.58	0.1151
Accuracy	0.36 (0.07)	0.15 (0.07)	0.21	1.98	0.0484
Pace	–0.02 (0.08)	0.27 (0.07)	–0.28	–2.67	0.0076
Movement quality	2.13 (0.07)	2.09 (0.07)	0.04	0.41	0.6846

Bonferroni correction was applied giving a significant *p*-value of 0.0026 (0.05/19) for identification of statistical differences; hand skill items exhibiting culture-DIF are in bold type.
[a] This item has been eliminated due to misfit.
SE: standard error.

had a difference greater than 0.5 logits between their adjusted and unadjusted ACHS ability measures. The Pearson's *r* correlation between the two sets of children's ability estimates exceeded 0.99.

DISCUSSION

The ACHS is a newly developed instrument that utilizes naturalistic observations to assess children's hand skills while performing a range of activities in real-life contexts. The activity items included in the ACHS have been reviewed by content experts to ensure cultural compatibility (10). The hand skill items also have been shown to possess the same meaning across different cultural groups; for example, "reaching for" a block represents identical meaning whether a child completes the action in different countries (20). Therefore the present study aimed at providing cross-cultural validity evidence for use of the ACHS with Australian and Taiwanese groups of children.

The results of this study supported the 6 levels of the ACHS rating scale with the pooled data from Australia and Taiwan. Most of the participating children in both groups exhibited valid hand skill response patterns when performing ACHS activities. With the exception of the Catching hand skill item, the ACHS items formed a unidimensional scale with this pooled sample. The findings are in agreement with earlier studies investigating these properties of the ACHS separately with each Australian and Taiwanese group (11, 13). However, the results identified a few ACHS activity items and hand skill items with DIF, indicating significantly varied item difficulties between Australian and Taiwanese participant groups.

Several potential factors may have contributed to the cross-cultural variability displayed in the ACHS with the pooled sample. For example, different cultural traditions could influence children's exposure to specific activities (14). In Australian contexts, ball-related activities (e.g. football, tennis, cricket, and rugby) are a significant part of children's lives. Taiwanese children have limited opportunities to take part in those outdoor ball activities since they usually live in tall buildings/apartments in densely populated urban areas. Instead, Taiwanese children are encouraged to participate more in pre-handwriting activities (e.g. using scissors and folding paper). These experiences may result in subtle differences in the difficulty levels of some ACHS activities that the Australian or Taiwanese children performed.

Different cultural traditions also may further influence parents' or teachers' demands and educational expectations placed on children and their hand skills (14, 15), thus resulting in hand skill items with cultural DIF. Because many academic skills that involve hand skill activities (e.g. handwriting, mathematics, computer, and literacy) are highly valued in Taiwanese society, Taiwanese children are usually expected or required to complete those activities as efficiently and accurately as possible. Parents also have high expectations for children to perform well academically within Taiwanese culture (15). In Australian learning contexts, however, children have more flexibility and parents/teachers tend not to put time pressure on their children completing activities (31). Hence, Australian and Taiwanese children may perform some hand skills at different levels of effectiveness due to different expectations and demands placed on them within individual cultural contexts.

Different group characteristics between Australian and Taiwanese samples also may have contributed to the discrepancies. There were more typically developing children in the Australian sample, while there was a larger clinical group in the Taiwanese sample. Moreover, the rater who conducted the assessments has a Taiwanese cultural background and may be less familiar with culturally appropriate ways that Australian children perform hand skill activities. These may have introduced systematic differences as one source of DIF and should be investigated in future studies.

Items displaying DIF are usually removed or adjusted in order to facilitate the pooling of data from different groups (29). However, it has been argued that removal of items exhibiting the highest levels of DIF may induce more DIF at the test score level (32). Therefore, we adopted DIF adjustment in the study; however, some problems remained. For example, one activity item and two hand skill items did not fit the model as the culture-specific items. It has also been argued that the adjustment/manipulation of splitting DIF data may be too complex for practical use in clinical settings (33), time-consuming, (34) and unnecessary (28, 29). Although we found some differences between adjusted and unadjusted children's ability estimates, the amount of difference was small and the correlation between the two sets of scores was extremely high. These findings indicate that the adjustment for culture-DIF in the ACHS seemed to have only a minor impact and may not be required in this cross-cultural design.

The findings of the study should be interpreted with caution since a limited number (e.g. fewer than 150) of participants were recruited from each cultural group. The generalizability of the findings is also limited to Australian and Taiwanese groups that included unequal proportion of children with and without disabilities. Future studies should include larger samples and comparable proportion of children, as well as include those from other countries (e.g. European or other Asian contexts) to examine the ACHS's cross-cultural validity. In addition, this study did not investigate non-uniform DIF, in which item-difficulty calibrations change across participants of different ability levels, and this issue should be investigated in future studies.

Previous studies found that ACHS inter-rater and test-retest reliability coefficients were lower than is desirable (i.e. $r > 0.90$), which may reflect the nature of an ecologically-based assessment (10). For example, contextual influences in natural situations may affect both raters and children, and raters have no opportunities to check children's performance repeatedly. Considering that limitation in reliability could hamper an assessment's validity, optimal rater training for using ACHS reliably should be established. Also, re-investigation of the ACHS's reliability in more controlled conditions or using a video-recorded approach may be warranted in future studies in order to determine reasons for its unsatisfactory reliability.

In conclusion, this study provides preliminary evidence to support the cross-cultural validity of the newly developed ACHS when used with groups of Australian and Taiwanese children. The Rasch analysis results indicate the appropriateness of ACHS's 6-level rating scale with this pooled sample. All but one hand skill items fitted the Rasch model, and there was a reasonable rate of person response patterns after removing a few misfitting children. Although DIF related to cultural difference was found in 7 activity items and 7 hand skill items, adjustment for cultural DIF had little practical impact on children's ability estimates and may not be required. Continued cross-cultural validation of the ACHS is needed to build cumulative evidence.

ACKNOWLEDGEMENTS

The authors would like to thank the content experts for their contributions in the development of this instrument. The support from the Victorian Department of Education and Early Childhood Development for providing ethics approval for the study and the families and children who participated in this study are also acknowledged. The first author was supported by the Endeavour International Postgraduate Research Scholarship, Monash Postgraduate Scholarship, Faculty Postgraduate Excellence Award, and Postgraduate Publication Award from Monash University when completing this study.

REFERENCES

1. Henderson A, Pehoski C, editors. Hand function in the child: foundations for remediation. 2nd edn. Philadelphia, PA: Mosby; 2006.
2. Kimmerle M, Mainwaring L, Borenstein M. The functional repertoire of the hand and its application to assessment. Am J Occup Ther 2003; 57: 489–498.
3. Chien CW, Brown T, McDonald R. A framework of children's hand skills for assessment and intervention. Child Care Health Dev 2009; 35: 873–884.
4. Folio MR, Fewell RR. Peabody Developmental Motor Scales: examiner's manual. 2nd edn. Austin, TX: PRO-ED; 2000.
5. Henderson SE, Sugden DA, Barnett AL. Movement Assessment Battery for Children-2. London: Pearson Assessment; 2007.
6. Ho ES, Clarke HM. Functional evaluation in children with congenital upper extremity malformations. Clin Plast Surg 2005; 32: 471–483.
7. World Health Organization. International Classification of Functioning, Disability, and Health. Geneva: World Health Organization; 2001.
8. Krumlinde-Sundholm L, Holmefur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. Dev Med Child Neurol 2007; 49: 259–264.
9. Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. Neurology 2004; 63: 1045–1052.
10. Chien CW, Brown T, McDonald R. Examining content validity and reliability of the Assessment of Children's Hand Skills (ACHS): a preliminary study. Am J Occup Ther 2010; 64: 756–767.
11. Chien CW. The Assessment of Children's Hand Skills (ACHS): theoretical framework, development, and validation [dissertation]. Melbourne, Australia: Monash University; 2010.
12. Fisher AG, Bryze K, Atchison BT. Naturalistic assessment of functional performance in school settings: reliability and validity of the School AMPS scales. J Outcome Meas 2000; 4: 491–512.
13. Chien CW, Brown T, McDonald R. Rasch analysis of the Assessment of Children's Hand Skills in children with and without disabilities. Res Dev Disabil 2011; 32: 253–261.
14. Case-Smith J. Hand skill development in the context of infant's play: birth to 2 years. In: Henderson A, Pehoski C, editors. Hand function in the child: foundations for remediation. 2nd ed. Philadelphia, PA: Mosby; 2006, p. 117–141.
15. Tseng MH. Development of pencil grip position in preschool children. Occup Ther J Res 1998; 18: 207–224.
16. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. J Rehabil Med 2003; 35: 105–115.
17. Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. Med Care 2004; 42: 37–48.
18. Linacre JM. Sample size and item calibration stability. Rasch Measure Trans 1994; 7: 328.
19. Linacre JM. FACETS: many faceted Rasch measurement computer program. Chicago, IL: MESA Press; 2009.
20. Stauffer LM, Fisher AG, Duran L. ADL performance of black Americans and white Americans on the assessment of motor and process skills. Am J Occup Ther 2000; 54: 607–613.
21. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates; 2007.
22. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas 2002; 3: 205–231.
23. Linacre JM. Detecting multidimensionality: which residual data-type works best? J Outcome Meas 1998; 2: 266–283.
24. Linacre JM. WINSTEPS computer program. Chicago, IL: MESA Press; 2008.
25. Scheuneman JD, Subhiyah RG. Evidence for the validity of a Rasch model technique for identifying differential item functioning. J Outcome Meas 1998; 2: 33–42.
26. Hagell P, Hoglund A, Reimer J, Eriksson B, Knutsson I, Widner H, et al. Measuring fatigue in Parkinson's disease: a psychometric study of two brief generic fatigue questionnaires. J Pain Symptom Manage 2006; 32: 420–432.
27. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ 1995; 310: 170.
28. Angoff WH. Perspectives on differential item functioning methodology. In: Holland PW, Wainer H, editors. Differential item functioning. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993, p. 3–23.
29. Crane PK, Hart DL, Gibbons LE, Cook KF. A 37-item shoulder functional status item pool had negligible differential item functioning. J Clin Epidemiol 2006; 59: 478–484.
30. Fischl C, Fisher AG. Development and Rasch analysis of the assessment of computer-related skills. Scand J Occup Ther 2007; 14: 126–135.
31. Veitch J, Bagley S, Ball K, Salmon J. Where do children usually play? A qualitative study of parents' perceptions of influences on children's active free-play. Health Place 2006; 12: 383–393.
32. Teresi JA. Different approaches to differential item functioning in health applications. Advantages, disadvantages and some neglected topics. Med Care 2006; 44: 152–170.
33. Elhan AH, Kutlay S, Kucukdeveci AA, Cotuk C, Oztürk G, Tesio L, et al. Psychometric properties of the Mini-Mental State Examination in patients with acquired brain injury in Turkey. J Rehabil Med 2005; 37: 306–311.
34. Hambleton RK. Good practices for identifying differential item functioning. Med Care 2006; 44: 182–188.