The Unilateral Below Elbow Test: a function test for children with unilateral congenital below elbow deficiency

Anita M Bagley* PhD, Shriners Hospital for Children Northern California;
Fred Molitor PhD, ETR Associates, Sacramento, California;
Lisa V Wagner OTR/L, Shriners Hospital for Children Greenville, Greenville, South Carolina;
Wendy Tomhave OTR/L, Shriners Hospital for Children Twin Cities, Minneapolis, Minnesota;
Michelle A James MD, Shriners Hospital for Children Northern California, Sacramento, California, USA.

**Correspondence to first author at* Motion Analysis Laboratory, Shriners Hospital for Children Northern California, 2425 Stockton Blvd., Sacramento, CA 95817, USA.
E-mail: abagley@shrinenet.org

The Unilateral Below Elbow Test (UBET) was developed to evaluate function in bimanual activities for both the prosthesis wearer and non-wearer. Nine tasks were chosen for each of four age-specific categories defined by development stages of hand function (2–4y, 5–7y, 8–10y, and 11–21y). Two scales, Completion of Task and Method of Use, were designed to rate performance. To measure reliability, four occupational therapists scored samples of videotaped UBET performances. For Completion of Task, an interval scale, agreement in scoring was measured with interclass correlation coefficients (ICC; *n*=9; five females, four males). For Method of Use, a nominal scale, chance-adjusted association was calculated with Cohen's kappa coefficients (interobserver *n*=198; 111 females, 87 males; intraobserver *n*=93; 56 females, 37 males). For Completion of Task, the average ICC was 0.87 for the prosthesis-on condition, and 0.85 for the prosthesis-off condition. ICCs exceeded 0.80 for eight out of nine tasks for the two older age groups, but for only five out of nine tasks in the younger age groups. Higher inter- and intraobserver kappa coefficients for Method of Use resulted when scoring children with their prostheses on versus off. The oldest age group had lower kappa values than the other three groups. The UBET is recommended for the functional evaluation of Completion of Task in children with unilateral congenital below elbow deficiency with and without their prostheses. Method of Use scoring can evaluate individuals for directed therapy interventions or prosthetic training.

See end of paper for list of abbreviations.

Unilateral congenital below elbow deficiency (UCBED) is the most common level of upper extremity transverse failure of formation (Giele et al. 2001) and the most amenable to prosthetic fitting (Sorbye 1980, Pruitt et al. 1997). Several different types of prostheses are available for children with UCBED, and pediatric prosthetic clinics typically provide prostheses for infants with this condition when they are able to sit independently, and prescribe prosthetic training (Singh and Varma 1988). However, little is known about how children with UCBED benefit from a prosthesis. Many of those fitted with a prosthesis eventually abandon it (Postema et al. 1999) or select a passive prosthesis without reporting diminution in function (Crandall and Tomhave 2002). The relationship between prosthesis wear and function is not well defined. A standardized tool that provides reproducible scores that measure how children with UCBED accomplish developmentally appropriate tasks with and without their prostheses is needed to understand the relationship between prosthesis use and function. Specifically, such a tool could determine whether prosthetic use improves function, how non-wearers use their residual limb during activities that generally require two hands, and whether different types of prostheses augment the performance of different activities.

Two tools currently exist that evaluate prosthetic function and have undergone reliability testing. The University of New Brunswick Test of Prosthetic Function is a well-established observational assessment of task performance that measures prosthetic skill and spontaneity but does not evaluate how children with UCBED function without a prosthesis (Sanderson and Scott 1985). The Prosthetic Upper Extremity Function Index (PUFI) is a questionnaire that addresses how children complete tasks with or without their prosthesis (Wright et al. 2001, 2003). It is not an observational evaluation of actual use nor is it designed to assess the child who does not wear a prosthesis.

For this study, children with UCBED were defined as wearers and non-wearers. Wearers were children who had used one or more than one prosthesis for any activity during the 6 months before the study. Non-wearers were children who either had never worn a prosthesis or had not used a prosthesis for any activity in the previous 6 months.

After recognizing the need to address the functional status of the wearer and non-wearer, the Shriners Hospitals for Children UCBED Study Group (see Acknowledgements) developed the Unilateral Below Elbow Test (UBET). This tool is intended to: (1) measure how children with UCBED use their prostheses to accomplish developmentally appropriate two-handed tasks; (2) measure how children with UCBED use their residual limb to accomplish the same tasks when they are not wearing a prosthesis; and (3) determine if the type of prosthesis used affects the child's ability to accomplish functional tasks. Before it could be used for these purposes, the validity of the UBET had to be established, including intra- and interobserver reliability (Trochim 2001). This paper describes the development of the UBET and reports reliability results.

Method

UBET DEVELOPMENT

Bimanual tasks that are important for activities of daily living were selected by the UCBED Study Group and additional experts in the field: the developers of the PUFI, several occupational therapists (OT), two prosthetists serving pediatric

populations, and a professional with statistical and tool development experience were used to carry out the study. Criteria for task selection included activities that require age-appropriate gross and fine motor skills, bimanual stabilization, manipulation, and application of resistive forces.

Nine tasks were chosen for each of the four age-specific categories defined on development stages of hand function: (1) 2 to 4 years; (2) 5 to 7 years; (3) 8 to 10 years; and (4) 11 to 21 years (see Appendix I). Children in the youngest age group develop skills of in-hand manipulation, controlled release, and appropriate grip force calibration (Henderson and Pehoski 1995). Children aged 5 to 7 years gain skill in intrinsic hand motions, but still do not fully demonstrate adult methods or speed. Children aged 8 to 10 years demonstrate greater efficiency in manipulation tasks and improved ability in asymmetric hand functions, linked to continuing neurological maturation (Henderson and Pehoski 1995). The oldest UBET age group should demonstrate fully mature and competent hand functions.

Two scoring scales, Completion of Task and Method of Use, were developed to assess function for each task. Completion of Task assesses level of function on a 5-point interval scale designed to distinguish ease of task completion and the quality of movement displayed by a child with UCBED when performing a task with or without the use of a prosthesis (Table I). Completion of Task is the primary score to be used to evaluate function because the essential assessment is whether or not the child can perform a specific task. This score can be used to compare performance between wearers and non-wearers, between prosthesis-on and -off conditions, and between prosthesis types. The scale for Completion of Task ranges from 0 to 4, with 4 representing completion of the task with no difficulty, and 0 representing inability to complete the task. To rate Completion of Task, the goal for success is specifically defined for each task; in the UBET administration instructions, the goal is underlined (Appendix II). For example, for Task 3 in the 2- to 4-year-old group, 'Put sock on foot' is defined as successfully completed if the sock is placed over the child's toes.

Method of Use assesses on a nominal scale how the child completes the task with and without a prosthesis (Table II). The purpose of this score is to assist the OT in developing patient-specific training sessions to learn the mechanics of a prosthesis. For Method of Use, the exact skill that should be evaluated is defined for each task. For 'Put sock on foot', Method of Use scoring is applied to how the sock is opened to initiate the task (Appendix II).

Standardized guidelines for test administration were developed, including order of task performance and instructions to the participant. The OT provides verbal instructions for all tasks, and demonstrates the Rolling Racer (a seated cart propelled with handlebars), bow and arrow, and telescope making tasks. To promote spontaneity, the evaluation is videotaped and Completion of Task and Method of Use scores are rated by the OT at a subsequent viewing. Performance of the UBET takes approximately 20 minutes. Children who wear a prosthesis perform the UBET twice, once with and once without their prosthesis, in randomized order. Full administration guidelines and a list of the standardized items to use for UBET administration may be found at www.shrinershq.org under Research.

Table I: Unilateral Below Elbow Test Completion of Task scores

Score	*Task*
4	Completes the task without difficulty. Movements are quick, smooth. Stability is maintained throughout task
3	Completes task with minimal difficulty. Movements are quick but slightly awkward. Stability is readily regained when lost
2	Completes task with moderate difficulty. Movements are slower and awkward
1	Completes task with maximal difficulty. Movements are very slow and awkward. Stability is frequently lost
0	Unable to complete task

Table II: Unilateral Below Elbow Test Method of Use scores

Coding	*Prosthesis-on*	*Prosthesis-off*
A	Active grasp of terminal device	Residual limb end manipulation and/or stabilization
P	Passive use of prosthetic forearm or terminal device	Forearm stabilization
E	Elbow or trunk grasp	Elbow or trunk grasp
N	No use of affected limb	No use of affected limb

INTER- AND INTRAOBSERVER RELIABILITY TESTING

Under a protocol approved by 11 Institutional Review Boards[1] as part of a multicenter clinical outcomes study of children with UCBED, the UBET and several questionnaires, including the PUFI and the Pediatric Outcomes Data Collection Instrument (PODCI), a musculoskeletal health questionnaire (American Academy of Orthopaedic Surgeons 2001), were administered to children at 10 Shriners Hospitals. Informed consent was obtained from all participants, or their parents as appropriate, based on age.

As part of the development of the UBET, four OTs thoroughly trained in the test scored the performance of seven wearers with their prosthesis on and off, and two non-wearers without a prosthesis. This sample was selected from different participating Shriners Hospitals, and included participants from all four age groups. Each OT independently recorded Completion of Task and interobserver reliability was measured. For Method of Use, OTs were paired to test interobserver reliability, and assigned sample sizes of 40 to 59 participants that included wearers and non-wearers from all four age categories. For intraobserver reliability testing, three OTs rescored Method of Use for a sample of children from their hospital. Sample sizes varied from 27 to 32 participants and again included wearers and non-wearers from all four age categories. Rescoring occurred at least 6 months after the initial scoring and OTs were blinded to their initial scoring.

STATISTICAL ANALYSIS

Completion of Task scoring was based on an interval scale, and interobserver reliability was calculated from the scores of all four OTs. Reliability was measured by calculating ICCs, which

[1]Because this was a multicenter study, approval from 11 local Institutional Review Boards was required.

provide a direct measure of agreement. Cohen's kappa coefficients were computed for the nominal-scaled Method of Use data. For Method of Use, interobserver reliability coefficients were calculated for the four defined pairs: (1) OT 1 versus OT 2; (2) OT 2 versus OT 3; (3) OT 3 versus OT 4; and (4) OT 4 vs OT 1, and intraobserver reliability coefficients were calculated for OTs 2, 3, and 4. Data from all OTs were combined for task-by-task analyses to obtain sufficient sample sizes.

Both percent agreement and Cohen's kappa statistics (Cohen 1960) were calculated for the Method of Use data. While percent agreement represents the degree to which the scores for paired OTs agree, the kappa coefficient adjusts for the degree of agreement for a coding scheme that could be due to chance. Thus, kappa values are often lower than the percent of agreement, and are considered a superior indication of reliability.

Twelve OTs participated and *rbo* values were calculated. But ICCs (reported in Table IIIa and b) were based just on the data from the four most experienced OTS.

Results

COMPLETION OF TASK

For the Completion of Task score, the average ICC was 0.87 for the prosthesis-on condition, and 0.85 for the prosthesis-off condition (Table IIIa). Coefficients exceeded 0.80 for six of the nine tasks in the prosthesis-on condition, and for seven of the nine tasks in the prosthesis-off condition. Average ICCs for the various age groups were 0.77 for the 2- to 4-year-olds, 0.79 for the 5- to 7-year-olds, 0.91 for the 8- to 10-year-olds, and 0.92 for the 11- to 21-year-olds (Table IIIb). Coefficients exceeded 0.80 for eight of nine tasks for the two older age groups but for only five of the nine tasks in the younger age groups.

METHOD OF USE

Method of Use interobserver reliability kappa coefficients ranged from 0.68 to 0.82 for scoring the prosthesis-on condition, and from 0.40 to 0.75 for scoring the prosthesis-off condition (Table IV). Analysis by age groups resulted in average coefficients of 0.69 for 2- to 4-year-olds, 0.81 for 5- to 7-year-olds, 0.78 for 8- to 10-year-olds, and 0.40 for 11- to 21-year-olds for the prosthesis-on condition (Table Va). Average interobserver coefficients for the prosthesis-off condition were 0.48 for 2- to 4-year-olds, 0.54 for 5- to 7-year-olds, 0.53 for 8- to 10-year-olds, and 0.38 for 11- to 21-year-olds (Table Vb).

Intraobserver reliability for Method of Use ranged from 0.70 to 0.85 for the prosthesis-on condition, and from 0.43 to 0.64 for the prosthesis-off condition (Table VI). Analysis by age group resulted in average intraobserver reliability coefficients of 0.68 for 2- to 4-year-olds, 0.69 for 5- to 7-year-olds, 0.73 for 8- to 10-year-olds, and 0.57 for 11- to 21-year-olds (Table VII).

Discussion

The UBET is an objective tool designed to assess the upper extremity function of children with UCBED, with and without a prosthesis. This study describes the development of the UBET and reports interobserver reliability results for Completion of Task scores, and inter- and intraobserver reliability results for Method of Use scores.

Data analysis shows good interobserver reliability for the Completion of Task scale. Reliability coefficients based on

Table IIIa: Interclass correlation coefficients for Completion of Task interobserver reliability[a]

Task	*Prosthesis-on (n=7)*	*Prosthesis-off (n=9)*
1	0.82	0.58
2	0.68	0.80
3	0.79	0.93
4	0.97	0.87
5	0.96	0.79
6	0.92	0.96
7	0.93	0.90
8	0.76	0.85
9	0.97	0.94
Average	0.87	0.85

Table IIIb: Interclass correlation coefficients for Completion of Task interobserver reliability for various age groups[a]

Task	*2–4y (n=4)*	*5–7y (n=4)*	*8–10y (n=5)*	*11–21y (n=3)*
1	0.33	0.88	0.93	1.00
2	0.67	1.00	0.89	1.00
3	0.89	0.56	0.97	0.92
4	0.95	0.80	0.98	0.97
5	0.69	0.73	0.98	0.98
6	0.83	0.86	0.98	0.44
7	0.77	0.97	0.53	0.99
8	1.00	0.65	0.91	0.94
9	0.82	0.62	0.99	0.99
Average	0.77	0.79	0.91	0.92

[a]Task numbers do not correspond to consistent tasks across age groups (see Appendix I). Grouped data are presented in columns.

Table IV: Interobserver percentage agreement and reliability coefficients for Method of Use by prosthesis condition

	OT 1 vs OT 2		*OT 2 vs OT 3*		*OT 3 vs OT 4*		*OT 4 vs OT 1*	
	% agree	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*
Prosthesis-on	79	0.71	92	0.82	81	0.68	88	0.81
	n=28		*n*=46		*n*=35		*n*=27	
Prosthesis-off	57	0.40	75	0.60	63	0.41	84	0.75
	n=41		*n*=59		*n*=52		*n*=40	

% agree, percentage of agreement between occupational therapists (OT).

age groups were limited by sample size (fewer observations per category) but still averaged between 0.77 and 0.92. Higher values (>0.90) were recorded for the two older age groups. Larger sample sizes would reduce the influence of outlier data and lead to improvement in these values. As interobserver reliability was not problematic for Completion of Task, intraobserver reliability analysis was not performed for this scale.

Good interobserver reliability for Method of Use was recorded for the prosthesis-on condition. When analyzed by age group, the oldest group (11–21y) demonstrated a relatively poor kappa value of 0.40. For this scale, both inter- and intraobserver reliability scores for the prosthesis-off condition were consistently lower than for the prosthesis-on condition.

Other tests or classification systems of pediatric function have reported inter- and intraobserver reliability data comparable to the results reported here. The Gross Motor Function Measure (Russell et al. 2002) assesses performance of gross motor tasks in children with cerebral palsy. The test consists of 88 tasks performed in the clinic setting. A 0- to 3-point scale is used to rate each task: (0) does not initiate; (1) initiates; (2) partially completes; and (3) completes. In addition to a total score, there are five dimension scores ranging from Lying & Rolling (lowest), to Walking, Running & Jumping (highest). Inter- and intraobserver reliability was measured from six OTs observing 12 children using test–retest data collected in a 2-week period. Intraobserver ICCs were 0.99 for the total score and all dimensions except Standing (0.92). Interobserver ICCs ranged from 0.87 to 0.99.

The Melbourne Assessment of Unilateral Upper Limb Function (Randall et al. 1999) is designed to score quality of motor function in reach, grasp, release, and manipulation for children 5 to 15 years of age with neurological impairment. Reliability testing was performed on 20 children by 16 OTs. Interobserver agreement for the 37 items ranged from 0.40 to 0.96, with an average kappa of 0.80. Coefficients for repeat scoring of items ranged from 0.11 to 1.00.

Bae et al. (2003) compared three functional classification systems for children with brachial plexus birth palsy: (1) the modified Mallet Classification; (2) Toronto Test Score; and (3) Hospital for Sick Children Active Movement Scale. Two physicians rated 80 children on two separate occasions. The average intraobserver kappa was 0.76 for the Mallet, 0.73 for the Toronto, and 0.85 for the Active Movement, with a minimum item value of 0.50. The average interobserver kappa was 0.78 for the Mallet, 0.51 for the Toronto, and 0.66 for the Active Movement, with a minimum item value of 0.21.

Table Va: Interobserver percentage agreement and reliability coefficients for Method of Use by prosthesis condition, age, and task[a]

Prosthesis-on	*2–4y (n=21)*		*5–7y (n=40)*		*8–10y (n=35)*		*11–21y (n=39)*	
	% agree	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*
Task 1	0.68	0.53	0.93	0.86	0.94	0.78	0.80	0.56
Task 2	0.89	0.81	0.90	0.75	0.83	0.67	0.56	0.37
Task 3	0.82	0.49	0.92	0.86	0.94	0.84	0.76	0.43
Task 4	0.76	0.64	0.80	0.73	0.97	0.93	0.68	0.41
Task 5	0.91	0.82	0.95	0.90	0.71	0.60	0.68	0.48
Task 6	0.68	0.46	0.83	0.73	0.97	0.94	0.56	0.30
Task 7	0.95	0.92	0.90	0.81	0.94	0.81	0.60	0.18
Task 8	0.84	0.77	0.93	0.86	0.91	0.72	0.60	0.41
Task 9	0.85	0.80	0.90	0.83	0.83	0.69	0.76	0.44
Average	0.82	0.69	0.89	0.81	0.90	0.78	0.67	0.40

[a]Task numbers do not correspond to consistent tasks across age groups (see Appendix I). Grouped data are presented in columns. % agree, percentage of agreement between occupational therapists.

Table Vb: Interobserver percentage agreement and reliability coefficients for Method of Use by prosthesis condition, age, and task[a]

Prosthesis-off	*2–4y (n=31)*		*5–7y (n=52)*		*8–10y (n=52)*		*11–21y (n=55)*	
	% agree	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*
Task 1	0.52	0.32	0.68	0.41	0.74	0.57	0.65	0.09
Task 2	0.77	0.52	0.90	0.64	0.80	0.65	0.70	0.56
Task 3	0.81	0.66	0.78	0.64	0.72	0.56	0.63	0.28
Task 4	0.83	0.64	0.61	0.44	0.54	0.31	0.74	0.54
Task 5	0.90	0.37	0.69	0.44	0.68	0.48	0.62	0.41
Task 6	0.68	0.55	0.75	0.65	0.90	0.78	0.52	0.24
Task 7	0.41	0.12	0.78	0.66	0.55	0.36	0.51	0.29
Task 8	0.83	0.57	0.66	0.39	0.72	0.47	0.80	0.72
Task 9	0.73	0.52	0.71	0.61	0.71	0.57	0.50	0.28
Average	0.72	0.48	0.73	0.54	0.71	0.53	0.63	0.38

[a]Task numbers do not correspond to consistent tasks across age groups (see Appendix I). Grouped data are presented in columns. % agree, percentage of agreement between occupational therapists.

Good inter- and intraobserver reliability for Method of Use scoring in the prosthesis-on condition was achieved (Tables IV and VI). For the prosthesis-off condition, kappa coefficients demonstrated wider variation, ranging from fair to good. Tables Va and Vb highlight the reliability issues with the prosthesis-off condition and with the oldest age group. Method of Use was easier to determine in the prosthesis-on condition because the action of the prosthetic terminal device is easier to discern than residual limb manipulation.

Differences in scoring Method of Use may be related to the many variations of grasp and stabilization that are routinely used in hand functions. For Method of Use scoring in the UBET, OTs were instructed to choose the code they felt best described how the child accomplished the task the majority of the time. The OTs had to select a single code per task although the child may have used more than one method of grasp during the activity. Also, evaluation of the motion of the affected limb from the videotape could be obscured by the position of the unaffected limb in tasks such as starting a zipper on a vest. As noted in Table VII, 'Start zipper on vest', which was Task 9 in the 8- to 10-year-old age group and Task 6 in the 11- to 21-year-old age group, demonstrated poor intraobserver reliability (0.29 and 0.24 respectively).

The oldest age group was the most difficult to score consistently for Method of Use. The interobserver data (Tables Va and Vb) show one task with poor reliability for this age group in each prosthesis condition: Task 7, tie shoelaces for the prosthesis-on condition (0.18); and Task 1, cut paper from a roll for the prosthesis-off condition (0.09). Three of the nine tasks for this age group demonstrated poor intraobserver reliability (Table VI): Task 3, secure wrapping paper with tape (0.12); Task 6, start a zipper (0.24); and Task 7, tie shoelaces (0.31). It is unclear why the shoelace task was not consistently scored in this age group. The same task in the 8- to 10-year-old group (Task 4) had an intraobserver reliability coefficient of 0.87.

The four OTs with the largest UCBED populations at their hospitals, and, therefore, with the most experience administering and scoring the UBET, had highest interobserver reliability in Method of Use. Therefore, the Method of Use reliability data represent results of practiced users of the UBET. Use of the videotapes to review and discuss cases can help less experienced observers develop consistency in Method of Use scoring. As the main purpose of this score is to direct an OT's treatment (prosthetic training) of a child, comparisons with other observers may be of limited clinical importance.

The UBET includes developmentally appropriate tasks using gross and fine motor skills, and activities of daily living for each age group. However, the sequence of tasks was not organized according to these parameters (i.e. Task 1 is not defined as fine motor for all age groups, Task 2 is not defined as gross motor for all age groups, etc.). Stratification in that manner may have helped to define characteristics of tasks across age groups that were difficult to score.

To date, the only other observational assessment tool designed for children with UCBED is the University of New Brunswick Test of Prosthetic Function (Sanderson and Scott 1985). Although this tool measures prosthetic function, it does not address how children with UCBED perform without a prosthesis. This is especially important because a large percentage of children with UCBED abandon their prostheses (Postema et al. 1999). In a recent survey, Shaperman et al. (2003) conclude that although objective data are needed to address prosthetic management, few clinics use standardized measurement tools of function.

Further validation of the UBET is underway, including tests of convergent validity with the PUFI and with the upper extremity physical function domain of the PODCI. Those two questionnaires were administered to all participants in the UCBED study, so correlation between observed function and parent or self-report of function will be performed for both prosthesis wearers and non-wearers. In addition, test–retest reliability of the UBET needs to be addressed in future work. Stability of performance may be assessed by having the same child perform the test twice over an interval of a few weeks; these data can provide further evidence of the clinical usefulness of the UBET.

Table VI: Intraobserver percentage agreement and reliability coefficients for Method of Use by prosthesis condition[a]

	OT 2		*OT 3*		*OT 4*	
	% agree	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*
Prosthesis-on	80	0.70	96	0.85	87	0.77
Prosthesis-off	78	0.64	78	0.63	63	0.43

[a]Occupational therapist (OT) 1 did not participate. % agree, percentage of agreement between OTs.

Table VII: Intraobserver percentage agreement and reliability coefficients for Method of Use by task and age group[a]

Task	*2–4y*		*5–7y*		*8–10y*		*11–21y*	
	% agree	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*	*% agree*	*kappa*
1	80	0.71	75	0.54	85	0.73	87	0.58
2	80	0.71	83	0.71	92	0.88	87	0.82
3	67	0.46	75	0.65	77	0.57	67	0.12
4	70	0.59	58	0.41	92	0.87	93	0.89
5	82	0.59	67	0.50	69	0.50	80	0.67
6	90	0.83	75	0.65	100	1.00	53	0.24
7	80	0.69	100	1.00	85	0.71	60	0.31
8	90	0.84	91	0.86	100	1.00	80	0.65
9	80	0.70	92	0.88	54	0.29	87	0.80
Overall	80	0.68	80	0.69	84	0.73	77	0.57

[a]Task numbers do not correspond to consistent tasks across age groups (see Appendix I). Grouped data are presented in columns. % agree, percentage of agreement between occupational therapists.

Conclusion

The UBET was designed to address a need for an objective functional assessment tool. New evaluation tools should be validated before they are accepted for clinical use. One important part of tool validation is the documentation of inter- and intraobserver reliability. The UBET has good inter-observer reliability globally for the Completion of Task score. This is the primary score to be used to judge ability and to compare groups such as wearers and non-wearers, or children with different types of prostheses. The Method of Use score demonstrates good inter- and intraobserver reliability for the prosthesis-on condition, and moderate inter- and intraobserver reliability for the prosthesis-off condition. This score should be used to direct patient-specific training and prosthetic prescription.

DOI: 10.1017/S0012162206001204

Accepted for publication 28th January 2006.

Acknowledgements

The development and validation of the UBET was part of a multicenter study of children with UCBED funded by Shriners Hospitals for Children Clinical Outcomes Studies Grant Program. The UBET was the product of the UCBED Study Group, which represents 10 Shriners Hospitals and includes: Katherine Brasington (Erie), Elaine Charest (Springfield), Leslie A Clawson (Northern California), Loray Dailey (St. Louis), Susan Duff (Philadelphia), Cheryl Hanley (Northern California), Becky Ligon (Houston), Cheryl Lutz (Philadelphia), Sharon McConnell (Erie), Kathleen Montpetit (Montreal), Joanna Patton (Los Angeles), and Joanne Shida (Los Angeles).We would also like to acknowledge Sheila Hubbard and Virginia Wright, developers of the PUFI for their helpful consultation, the UCBED grant participants for their time and commitment to the promotion of better care for children with UCBED, and finally the participants and their parents for their willingness to take part in this study.

References

American Academy of Orthopaedic Surgeons. (2001) *AAOS Normative Data Study and Outcomes Instruments.* Rosemont: AAOS

Bae DS, Waters PM, Zurakowski D. (2003) Reliability of three classification systems measuring active motion in brachial plexus birth palsy. *J Bone Joint Surg Am* **85-A:** 1733–1738.

Cohen J. (1960) A coefficient of agreement for nominal scales. *Educ Psych Meas* **20:** 37–46.

Crandall RC, Tomhave W. (2002) Pediatric unilateral below-elbow amputees: retrospective analysis of 34 patients given multiple prosthetic options. *J Pediatr Orthop* **22:** 380–383.

Giele H, Giele C, Bower C, Allison M. (2001) The incidence and epidemiology of congenital upper limb anomalies: a total population study. *J Hand Surg Am* **26:** 628–634.

Henderson A, Pehoski C. (1995) *Hand Function in the Child: Foundations for Remediation.* St Louis: Mosby.

Postema K, Van der Donk V, Van Limbeek J, Rijken RA, Poelma MJ. (1999) Prosthesis rejection in children with a unilateral congenital arm defect. *Clin Rehabil* **13:** 243–249.

Pruitt SD, Varni JW, Seid M, Setoguchi Y. (1997) Prosthesis satisfaction outcome measurement in pediatric limb deficiency. *Arch Phys Med Rehabil* **78:** 750–754.

Randall MJ, Johnson LM, Reddihough DS. (1999) *The Melbourne Assessment of Unilateral Upper Limb Function. Test Administration Manual.* Royal Children's Hospital: Melbourne.

Russell DJ, Rosenbaum PL, Avery LM, Lane M. (2002) *Gross Motor Function Measure (GMFM-66 & GMFM-88) User's Manual. Clinics in Developmental Medicine No. 159.* London: Mac Keith Press

Sanderson ER, Scott RN. (1985) *UNB Test of Prosthetics Function.* New Brunswick: Bio-Engineering Institute.

Shaperman J, Landsberger SE, Setoguchi Y. (2003) Early upper limb prosthesis fitting: when and what do we fit. *J Prosthet Orthot* **15:** 11–17.

Singh U, Varma SK. (1988) Rehabilitative management of upper limb amputees. *Indian J Med Res* **88:** 99–106.

Sorbye R. (1980) Myoelectric prosthetic fitting in young children. *Clin Orthop Relat Res* **148:** 34–40.

Trochim WMK. (2001) *The Research Methods Knowledge Base.* Cincinnati: Atomic Dog Publishing. p 96–97.

Wright FV, Hubbard S, Jutai J, Naumann S. (2001) The Prosthetic Upper Extremity Functional Index: development and reliability testing of a new functional status questionnaire for children who use upper extremity prostheses. *J Hand Ther* **14:** 91–104.

Wright FV, Hubbard S, Naumann S, Jutai J. (2003) Evaluation of the validity of the prosthetic upper extremity functional index for children. *Arch Phys Med Rehabil* **84:** 518–527.

List of abbreviations

OT	Occupational therapist
PODCI	Pediatric Outcomes Data Collection Instrument
PUFI	Prosthetic Upper Extremity Function Index
UBET	Unilateral Below Elbow Test
UCBED	Unilateral congenital below elbow deficiency

Appendix I: Summary of Unilateral Below Elbow Test (UBET) tasks

	2–4y	*5–7y*	*8–10y*	*11–21y*
Task 1	Take Play-Doh out of a plastic bag	Cut paper circle from construction paper	Swing a bat	Cut paper from a roll to wrap a videotape
Task 2	Bang cymbals together	Remove cap from felt tip marker	Wind string onto yo-yo	Tear a piece of tape
Task 3	Put sock on foot	Sharpen pencil	Open a Band-Aid	Secure wrap on videotape with tape
Task 4	Thread beads	Do up buttons on vest	Tie shoelaces in a bow	Cut putty on plate with knife and fork
Task 5	Open a jar of bubbles	Tie shoelaces in a knot	Do up buttons on a shirt	Open a three ring binder
Task 6	Ride on a Rolling Racer	Turn kaleidoscope	Make a telescope with paper and a rubber band	Start zipper on vest
Task 7	Open drawstring bag and dump LEGO DUPLOs out	Separate LEGOs	Place glove on unaffected hand	Tie shoelaces in a bow
Task 8	Separate LEGO DUPLOs	Use bow and arrow	Draw a line with a ruler	Do up buttons on a shirt
Task 9	Open a box of crayons and remove one	Ride on Rolling Racer	Start zipper on vest	Use dust pan and small broom

Full UBET Administration Instructions and a list of standardized items to be purchased are located on www.shrinershq.org under Research.

Appendix II: Sample Unilateral Below Elbow Test tasks

For 2–4y

1. Take Play-Doh out of a plastic bag. (Use the amount of Play-Doh from a full can. Make sure the Play-Doh is fresh. Flatten the Play-Doh and place it in the plastic bag with the top folded over so the Play-Doh will not simply fall out when the bag is turned over)

Score how plastic bag is stabilized

	Prosthesis-on	*Prosthesis-off*
Completion of Task score		
Method of Use score		
Demonstration		

2. Bang cymbals together

Score how cymbal is held or stabilized

	Prosthesis-on	*Prosthesis-off*
Completion of Task score		
Method of Use score		
Demonstration		

3. Put sock on foot. (Provide tube sock. Child should get sock over toes; okay if foot is on floor)

Score how sock is opened to initiate task

	Prosthesis-on	*Prosthesis-off*
Completion of Task score		
Method of Use score		
Demonstration		

4. Thread beads. (Child should string at least one 1 inch square bead)

Score how items are stabilized

	Prosthesis-on	*Prosthesis-off*
Completion of Task score		
Method of Use score		
Demonstration		