The development and reliability of the Cantonese version of the Allen Cognitive Level Screen

CHRISTOPHER LF CHAN Stewards' Limited, Hong Kong
SHARON RAY Tufts University, Boston, MA, USA
SCOTT TRUDEAU Tufts University, Boston, MA, USA

ABSTRACT: The Allen Cognitive Level Screen is a quick screening test to assess the cognitive functions of people with cognitive impairments or psychiatric disabilities. The purposes of the study were to translate the Allen Cognitive Level Screen into Cantonese and to gather evidence of the reliability of the translated version. Translation was performed by three bilingual occupational therapists. A panel of another five bilingual occupational therapists verified the accuracy of translation. Thirty randomly selected Cantonese-speaking healthcare workers performed the Cantonese version of the Allen Cognitive Level Screen. Results suggested that the test seemed to be accurately translated. Inter-rater reliability and the test-retest coefficient of the Cantonese version of the Allen Cognitive Level Screen were 0.98 and 0.73 (test-retest interval = 28.3 days) respectively. Future research should be directed towards further exploring the psychometric properties and clinical application of the Cantonese version of Allen Cognitive Level Screen.

Key words: Allen Cognitive Level Screen Test, cognitive assessment, cross-cultural measurement, reliability.

Introduction

The Allen Cognitive Level Test (ACL) is widely used by occupational therapists in North America, Israel and Australia (Keller and Hayes, 1998). It is built on the theoretical framework of the Cognitive Disability Model and is designed to estimate the cognitive level of persons with psychiatric illnesses or cognitive impairments (Allen et al., 1992). Each cognitive level reflects a person's abilities to perform familiar activities, the type of assistance needed to complete tasks safely, and the rehabilitation necessary to relearn tasks (Allen et al., 1992). The validity and reliability of the ACL have been shown in several studies (Mayer, 1988; David and Riley, 1990; Allen et al., 1992; Velligan
et al., 1995). Owing to its clinical usefulness and established psychometric properties, occupational therapists from different countries have translated the ACL into their own language – for example, French, and Japanese – and applied the translated version of the ACL to different cultural groups (Cusick and Harai, 1991). The purposes of this study were to translate the ACL into Cantonese and to examine the reliability of the Cantonese version of the ACL.

**Literature review**

The Cognitive Disability Model

The Cognitive Disability Model was developed by Claudia Allen and was based on her intensive observations and empirical research in the field of psychiatry (Allen, 1985). According to the model, cognitive disability is defined as ‘a restriction in voluntary motor action originating in the physical and chemical structures of the brain and producing observable limitation in routine task behavior’ (Allen, 1985: 31). A hierarchy of six cognitive levels is proposed in the model to describe progressively severe cognitive impairment (that is, 1 = severe cognitive impairment, 6 = normal cognition). Attention to sensory cues, motor actions and sensorimotor association are the main attributes of each cognitive level and are organized into an information processing system. The input represents sensory cues, the throughput is sensorimotor association and the output involves motor actions. Based on this system, actions are observable activity performance, which are elicited by sensory cues and guided by sensorimotor association (Allen, 1985; Allen, 1987; Levy, 1993). An overview of each cognitive level is as follows:

**Level 1 (Automatic action):** The person is conscious and cooperates with vital tasks through the use of sharp commands (Allen, 1987; David and Riley, 1990). According to Allen (1987), arousal is the term used to describe the basic ability of the person at this level to respond to external cues. A change in level of arousal is usually transient and must be initiated by external stimuli. The person’s behaviours are mostly reflexive or habitual. Training is not possible and the person requires 24-hour nursing care (Allen, 1985; Allen, 1987; David and Riley, 1990).

**Level 2 (Postural action):** The person’s behaviours often relate to a feeling of comfort or discomfort. People at level 2 are not aware of the external environment only of the movement of their body. Imitation of gross body movements is possible through demonstration and verbal cues but their spontaneous postural actions seem to be unique (Allen, 1985; Allen, 1987; David and Riley, 1990). Discernment is the term used to describe the limited ability of the person at this level to perceive the external environment and to recognize the external as separate from the self. Twenty-four-hour nursing care is usually necessary for an individual at cognitive level 2 (Allen, 1987).
Level 3 (Manual actions): The spontaneous actions observed in this level are manual actions. Actions are often repetitive, inappropriate, unpredictable or destructive to others, and are not goal-directed. Attention shifts from the internal self to the external environment but is restricted to tactile cues. Constant guidance is required for the person to complete a task successfully (Allen, 1985; Allen, 1987; David and Riley, 1990). Long-term repetitive training is the term used to describe the ability of people at this level to acquire a limited awareness of routine activities that are monitored by a caregiver (Allen, 1987). Twenty-four-hour supervision is usually necessary (Allen, 1987).

Level 4 (Goal-directed actions): The person's behaviours observed in this level are goal-directed and his or her attention is captured by visual cues as well as tactile cues. Reliance on visual cues may pose a danger to the person because electricity, heat or chemical stimuli do not gain his or her attention. Daily on-site supervision is necessary to avoid accidents that are a result of lack of attention to hazards (Allen, 1985; Allen, 1987; David and Riley, 1990). Allen (1987) used the term training to describe the limited capacity of a person at level 4 for learning, which is situation-specific; generalization to other situations is not feasible (David and Riley, 1990).

Level 5 (Exploratory actions): The person at this level explores the effect of his or her actions on the physical properties of the environment. According to Allen (1987), learning by doing or being shown what to do is the term used to describe overt trial or error processes of the persons functioning at this level, and generalization to other activities and environments is possible. People often fail to understand explanations that use symbolic concepts and do not anticipate future events. Situations that require planning, organization and deductive reasoning are usually not handled effectively. Supervisory assistance to anticipate safety hazards and planning is required (Allen, 1985; Allen, 1987; David and Riley, 1990).

Level 6 (Planned actions): This is the theoretical functional level of a normal adult, which indicates absence of disability. People at this level are able to use symbolic cues to formulate plans and to consider alternatives while engaged in an activity (David and Riley, 1990). Future events are anticipated, behaviour is organized and satisfactory results are produced (Allen, 1987).

These levels are used by occupational therapists as guidelines to predict the patient's ability to do familiar activities, to learn new activities, and for determining the amount of assistance needed from caregivers (Allen et al., 1992). The ACL is a practical screening test that is based on the Cognitive Disability Model for assessing the cognitive level of an individual.

The Allen Cognitive Level Test

The Allen Cognitive Level Test is an individualized criterion-referenced tool designed for assessing the cognitive level of clients with psychiatric illness or
cognitive impairment. There are six versions of the ACL, including (a) the Allen Cognitive Level Test – Original (ACL – O) (Allen, 1985); (b) the Allen Cognitive Level Test – Expanded (ACL – E) (Allen et al., 1992); (c) the Allen Cognitive Level Test – Problem Solving (ACL – PS) (Allen et al., 1992); (d) the Allen Cognitive Level Test – 90 (ACL – 90) (Henry et al., 1998); (e) the Large Allen Cognitive Level Test (LACL) (Allen et al., 1992); and (f) the Allen Cognitive Level Screen (ACLS) (Allen, 1996). The ACL – 90 is the version most commonly used by occupational therapists in the field of psychiatry and geriatrics (Allen et al., 1992; Velligan et al., 1998). However, the ACLS is the most refined version of the test (Allen, 1996).

The ACLS requires an individual to perform a series of sensorimotor tasks – specifically, three increasingly complex leather stitches. The individual is instructed to observe and replicate what the examiner does. The examiner then makes mistakes on purpose and asks the patient to point out the mistake and fix it. Finally, the individual is required to figure out the last stitch from a completed model. The scoring is dependent on the absence of errors in the stitching, the ability to correct errors and the ability to complete more complex stitches without assistance. Scores are rated on a 25-point scale, ranging from 3.0 to 5.8, with higher scores indicating higher cognitive levels. The score on the ACLS estimates a subject’s ability to follow directions, to solve problems and to learn new information (Allen, 1996).

Various versions of the ACL have shown excellent inter-rater reliability, with the correlation coefficient ranging from 0.92 to 0.99 (Moore, 1978; Penny et al., 1995; Henry et al., 1998). Moderate test-retest reliability of the ACL – O with a correlation coefficient of 0.75 was reported in a study by Newman (1987). ACL scores were found to correlate positively with cognitive and neurocognitive measures that assess fluid ability and executive functions respectively (Mayer, 1988; David and Riley, 1990; Velligan et al., 1998; Secrest et al., 2000). Significant and positive relationships between the ACL scores and measures of adaptive functioning were reported (Heying, 1985; Velligan et al., 1995; Henry et al., 1998; Keller and Hayes, 1998). These studies support the use of the ACL as a quick measure of a person’s cognitive and functional abilities.

The ACL is based on the assumption that the ‘stitch is a familiar, repetitive manual action for most American adults’ (Allen, 1985: 108). Examining the relevancy of the ACL to other ethnic groups is very important to determine the applicability of the test to these groups. Velligan and colleagues (1995) reported there was no main effect for ethnicity on the ACL scores in samples of non-Hispanic white, Mexican-American and African-American people. Authors reasoned that non-verbal problem abilities in the ACL might be less culturally biased than verbal tests. A review of the literature about the use of the ACL in Israel suggested the ACL seemed to be applicable to Israeli culture (Katz and Heiman, 1990). In addition, the Japanese ACL was reported to differentiate between disabled and non-disabled groups, suggesting that the
activity used in the ACL was suitable for the Japanese cultural context (Cusick and Harai, 1991).

Translation and validation
Developing an instrument that is based on an existing instrument from another culture generally involves translation and validation. Proper procedures in both translation and validation ensure that the new instrument is accurately translated, reliable and valid (Chapman and Carter, 1979). In translation, researchers often use backward translation, meaning that a second person who speaks both languages translates the translator's version back to the original language to see how close the meaning is to the original language (Weiss, 1997). Although backward translation is the most common and recommended procedure for verifying a translation, finding competent translators is often difficult (Chapman and Carter, 1979). Using other techniques instead of backward translation for verifying translation is not unusual – for example, the committee approach and the pretest procedure (Campbell et al., 1970). In the committee approach a group of bilingual individuals translate the source to the target language. The mistakes of one member of the group can be caught by others in the committee. The pretest procedure refers to field testing to ensure the subjects will understand all the questions (Brislin, 1970).

Following the translation, the next step in the validation process is to examine the reliability and validity of the translated instrument. The importance of using reliable and valid instruments in the practice of occupational therapy is recognized (Benson and Clark, 1982; Ottenbacher, 1987). During the validation of an instrument, its reliability is usually tested before its validity (Gregory, 1999). Test-retest and inter-rater reliability are the most commonly used methods used by researchers to substantiate the reliability of an instrument. Test-retest reliability refers to the stability of the test scores across time (Portney and Watkins, 1993). The most straightforward method to determine the test-retest reliability of an instrument is to administer the test twice to the same group of representative subjects (Gregory, 1999). Inter-rater reliability measures the variation between two or more raters who test the same group of subjects (Portney and Watkins, 1993).

The purposes of this study were to translate the ACLS into Cantonese and to examine the reliability of the Cantonese version of the ACLS. The study provided a foundation for the future validation of the Cantonese version of the ACLS.

Method
The study was divided into two main stages: translation of the ACLS, and measurement of the reliability of the Cantonese version of the ACLS. The
translation involved four steps that, using methodology proposed by Campbell et al. (1970), examined the accuracy of the translation.

Translation

Panels composed of bilingual occupational therapists and Cantonese-speaking healthcare workers were used at different steps of the study. Bilingual occupational therapists, who had extensive English and Chinese language training and experience in evaluating or treating patients with psychiatric illnesses, were selected to translate the ACLS or to evaluate the accuracy of the translation. Cantonese-speaking healthcare workers were included to evaluate the understandability of the translation.

Step 1: The ACLS has not been copyrighted, so its translation was initiated. Three bilingual occupational therapists, who have worked as occupational therapists in Hong Kong and the United States, were selected to translate the directions in the manual of the ACLS (Allen, 1996). The therapists received one hour of training before the translation. The training included an introduction to the ACLS, and information about the purpose of the study and translation techniques. The emphasis of the translation was to look at the conceptual rather than the literal meaning (Chang et al., 1999). The Cantonese version of the ACLS 1 was formulated.

Step 2: In this step, the Cantonese version of the ACLS 1 was evaluated in terms of the accuracy and equivalency of the translation. A convenience sample of five bilingual occupational therapists working in psychiatric day hospitals or psychiatric inpatient hospitals was selected. The selected therapists were contacted by phone individually to explain the purpose of the study and introduce the ACLS. A questionnaire, a copy of the original test manual and a return envelope were mailed to each therapist. A questionnaire comprising 42 questions was constructed for the therapists to evaluate the degree of accuracy and equivalency in the translation. Each question included a sentence from the original version of the ACLS’s instructions and the corresponding Cantonese version of the test instructions. Also, each question consisted of a closed-end portion and an open-end portion. The closed-end portion required the therapists to assign a rating on a four-point Likert scale whereas the open-end portion was used to solicit suggestions and comments from the therapists if they selected the rating 1 or 2. An example of a question presented is as follows:

The original version: ‘Have you ever done anything like this before?’

The Cantonese version: Is the sentence translated with the same meaning when compared with the original version?

1 Strongly disagree 2 Disagree 3 Agree 4 Strongly agree

If the rating is 1 or 2, please make suggestion:
Step 3: The Cantonese version of the ACLS 1 was revised by the translators after receiving all of the questionnaires. Each therapist was then mailed the Cantonese version of the ACLS 2 together with the previous questionnaires. The therapists then met to discuss the modified items in the Cantonese version of the ACLS until they reached a consensus on the modified items. The Cantonese version of the ACLS 3 was then formulated.

Step 4: The Cantonese version of the ACLS 3 was pretested with four Cantonese-speaking healthcare workers to examine the clarity and understandability of the translated test. The ethnicity of all workers was Chinese and their primary language was Cantonese. The Cantonese version of the ACLS 3 was administered to each worker after obtaining written consent. The workers' performances on the test were videotaped. Following the test, structured interviews comprising four questions relating to the understandability, clarity, conciseness and the need for further modification of the test instructions were conducted with the workers. A gain, each question consisted of a closed-end portion and an open-end portion. The closed-end portion required workers to assign a rating on a four-point Likert scale and the open-end portion was used to identify the specific part of test instruction that required further modification if the rating of 1 or 2 was selected. An example of a question in the structured interview was illustrated as follows:

Are the test instructions easily understood?

1 2 3 4
Strongly disagree Disagree Agree Strongly agree

If the rating is 1 or 2, please elaborate which part of test instructions you don’t understand. If the worker has difficulty in recalling, play back the videotape.

Because all ratings in the interview fell into category 3 or 4, the Cantonese version of the ACLS 3 was not revised further. The Cantonese version of the ACLS was finalized at this point.

Reliability study

Fifty-four employees in a skilled nursing facility met the selection criteria as Cantonese-speaking healthcare workers. The selection criteria were that workers:

• were Chinese;
• had Cantonese as their primary language;
• had worked in the facility for more than three months;
• were 18 to 60 years old;
• were able to read a newspaper with or without glasses;
• voluntarily gave their consent to participate in the study.
Thirty workers were randomly selected from a group of 54 to participate in the reliability study of the Cantonese version of the ACLS. Twenty-four (80%) out of 30 participants were female and six participants were male. The characteristics of the workers are summarized in Table 1.

The administration of the Cantonese version of the ACLS took place in a room in the skilled nursing facility that was free of distraction. Each participant's performance on the test was videotaped. The video recording focused only on the participant's hand to ensure confidentiality. The Cantonese version of the ACLS was readministered to all participants in the same location one month later. Again, the performance of the participants on the retest was videotaped.

Three raters, who were bilingual occupational therapists and who had experience of working with psychiatric and geriatric clients, were selected. The raters were randomly assigned as first, second and third rater. The raters were given a one-hour training session on the scoring of the Cantonese version of the ACLS. The first rater scored all performances in the first trial of test and retest by reviewing the videotapes while the second and third raters scored all performances only on the first trial of the test. To minimize the influence on each other during scoring, all raters scored the performances independently.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cantonese healthcare workers (%) (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20 (n=6)</td>
</tr>
<tr>
<td>Female</td>
<td>80 (n=24)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>21–30</td>
<td>20.0 (n=6)</td>
</tr>
<tr>
<td>31–40</td>
<td>23.3 (n=7)</td>
</tr>
<tr>
<td>41–50</td>
<td>33.4 (n=10)</td>
</tr>
<tr>
<td>51–60</td>
<td>23.3 (n=7)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Middle school</td>
<td>20.0 (n=6)</td>
</tr>
<tr>
<td>High school</td>
<td>30.0 (n=9)</td>
</tr>
<tr>
<td>Associate's degree</td>
<td>10.0 (n=3)</td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>36.7 (n=11)</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>3.3 (n=1)</td>
</tr>
<tr>
<td>Nature of work</td>
<td></td>
</tr>
<tr>
<td>Nursing</td>
<td>56.7 (n=17)</td>
</tr>
<tr>
<td>Recreation/activity</td>
<td>13.3 (n=4)</td>
</tr>
<tr>
<td>Administration</td>
<td>10 (n=3)</td>
</tr>
<tr>
<td>Maintenance</td>
<td>10 (n=3)</td>
</tr>
<tr>
<td>Other (rehabilitation and catering)</td>
<td>10 (n=3)</td>
</tr>
</tbody>
</table>
Test-retest and inter-rater reliability were determined by using the intra-class correlation coefficient (ICC). Among different models of ICC, model three with single measure was selected (Portney and Watkins, 1993).

**Results**

**Translation of the Allen Cognitive Level Screen**

Five panel members completed and returned the questionnaires, which were constructed to evaluate the accuracy and equivalence of the translation. Of the 42 questions in the questionnaires, 21 had ratings of 3 or 4, indicating no need for revision. However, the remaining 21 questions had ratings of 1 or 2, indicating a need for a revised translation. Ratings on the four-point Likert scale were regrouped and dichotomized into inaccurate translation (ratings of 1 or 2) and accurate translation (ratings of 3 or 4). The percentage of agreement between the panel members in the dichotomized scale ranged from 73.8% to 95.2%, with a mean of 80%.

There was no clear pattern for the types of sentence that needed a revised translation, as reflected in the ratings. Revision of the instructions was based on the comments from the panel members, leading to the Cantonese version of the ACLS 2. Additional minor revision was made on four test sentences after discussion among the panel members and the researcher.

**Pilot testing of the Cantonese version of the Allen Cognitive Level Screen**

Four participants completed the Cantonese version of the ACLS 3 and were asked to assign a rating for the four questions on a four-point Likert scale along four dimensions. No additional revision was performed because all ratings assigned by the participants fell into categories 3 or 4. The Cantonese version of the ACLS was finalized at this point.

**Measuring the reliability of the Cantonese version of the Allen Cognitive Level Screen**

Scores in the first trial of the test ranged from 5.4 to 5.8. The results of the test scores are summarized in Table 2. The inter-rater reliability coefficients for the Cantonese version of the ACLS, as determined by the intra-class correlation, was 0.98. The percentage agreements between three raters ranged from 96.7% to 100%.

Test scores on the retest also ranged from 5.4 to 5.8. The results of the test scores are summarized in Table 3. The average interval between test and retest was 28.3 days, with a range from 25 to 35 days. The test-retest reliability for the Cantonese version of the ACLS, as determined by the intra-class correlation, was 0.98.
correlation, was 0.73. The percentage agreement between the test scores on the first trial of test and retest was 80%.

Discussion

The aim of this study was to develop a Cantonese cognitive instrument for people with a psychiatric or cognitive disability that would accurately translate from an existing instrument and that would demonstrate good reliability over time and consistency among raters. This objective was reached.

Although the most commonly used translation technique, backward translation, was not used in the study, several techniques were used to ensure the accuracy of translation, including the committee approach, panel review and discussion, and pilot testing. As the results show, minor revision of half of the test instructions and the high agreement between the five panel members about the accuracy of translation showed that the quality of the initial translation of the test was satisfactory. The revised translation was tested further by checking it through a panel discussion and pilot testing. Minor revision of four out of 42 test instructions was recommended after a panel discussion, but no additional revision was performed after pilot testing. Based on the above argument, it is concluded that the translation of the Cantonese version of the ACLS seemed to be accurate.

The translated test still needs further validation before being applied clinically. The process of validation can facilitate a better understanding of the test instrument in term of consistency and the meaningfulness of the data. During the validation, an instrument’s reliability is usually tested before its validity (Gregory, 1999). In this study, the intra-class correlation was used to compute the coefficient. Unlike Pearson and Spearman’s correlation, the intra-class

| TABLE 2: Descriptive statistics of the participants' test scores rated by the three raters (n=30) |
|---------------------------------|------------|------|------|--------------|------|
|                                 | Mean       | Mode | Median| Standard deviation | Range |
| Rater 1                         | 5.65       | 5.6  | 5.6   | 0.15              | 5.4-5.8|
| Rater 2                         | 5.64       | 5.6  | 5.6   | 0.15              | 5.4-5.8|
| Rater 3                         | 5.64       | 5.6  | 5.6   | 0.15              | 5.4-5.8|

| TABLE 3: Descriptive statistics of the participants' test scores on the first trial of test and retest (n=30) |
|---------------------------------|------------|------|------|--------------|------|
|                                 | Mean       | Mode | Median| Standard deviation | Range |
| Test                            | 5.65       | 5.6  | 5.6   | 0.15              | 5.4-5.8|
| Retest                          | 5.70       | 5.8  | 5.8   | 0.14              | 5.4-5.8|
correlation reflects both the degree of consistency and agreement of ratings (Portney and Watkins, 1993). Some authors suggest that an intra-class correlation coefficient of 0.8 or higher is acceptable and shows good reliability (Gelinas et al., 1999).

Comparing the inter-rater reliability coefficient of the Cantonese version of the ACLS against the acceptable coefficient suggested by Gelinas et al. (1999) shows a high value for reliability. This finding is consistent with the result of the study of inter-rater reliability of the English version of the ACLS (Henry et al., 1999). This indicates that the scoring criteria of the ACLS are clear and well written. However, the coefficient of the Cantonese version of the ACLS is slightly below 1, which may result from lack of variability of the test scores or error by raters.

Reliability is estimated by the statistical concept of variance and can be expressed as a ratio of True variance/(True variance + Error variance) (Gregory, 1999). If the sample is homogeneous, the true variance will be reduced, which in turn reduces the reliability coefficient (Portney and Watkins, 1993). However, according to the between-subject variance in the analysis of variance, the subject variance was statistically significant (F=137.6, p<0.000), which implied that there was a great variability of test scores between subjects. Lack of variability of the test scores did not exert an influence in the value of the inter-rater reliability coefficient. Conversely, the percentage of agreement between the raters showed that raters 2 and 3 were perfectly agreed. The percentage of agreement for rater 1 was 96.7% in relation to raters 2 and 3 as there was a single disagreement in the test score of rater 1. This reduced the inter-rater reliability coefficient of the Cantonese version of the ACLS.

The test–retest reliability coefficient of the Cantonese version of the ACLS was slightly below the recommended value of an acceptable coefficient as suggested by Gelinas et al. (1999), which indicated that the test's stability over time is fair. The finding is also consistent with the result of study conducted by Newman (1987), who reported that the test–retest reliability coefficient of the ACL–O was 0.75. Lack of variability of the test scores or disagreement of the test scores in test and retest may account for the fair test–retest reliability.

A lack of variability of the test scores secondary to the homogeneous sample may reduce the test–retest reliability coefficient. In this case, the between-subject variance in the test scores was statistically significant (F=6.34, p<0.000), which indicated that there was a great variability in the test scores. On the other hand, according to the classical theory of measurement, the major source of error in test–retest reliability is changes over time (Gregory, 1999). The change may result from a real change in the traits of those persons, unstable traits, testing effects, test–retest interval, motivation factors of subjects, or rater bias (Portney and Watkins, 1993; Gregory, 1999). Because all participants in the reliability study are considered to have stable
cognitive ability, a change in cognitive ability during the test and retest period may not be possible. Therefore, the change of scores in the test–retest may not be due to any real change in traits or unstable traits in those persons. The motivation of the participants, such as improving their scores in the retest, may influence the test–retest reliability coefficient. Because all participants in the reliability study were blind to their test scores, it is not likely that the change of test scores in test and retest resulted from the motivation factor. Rater bias refers to the rater remembering the first test score when that rater is involved in both test and retest. However, in this study, several measures were taken to minimize the rater bias – for example, raters were blind to the identity of the participants and also rated the scores by reviewing videotapes that showed only the hands of participants. Rater bias did not influence the value of the coefficient. Learning occurring during repeated measures is very common (Portney and Watkins, 1993). The raw data showed that six pairs of ratings did not have retest scores that were higher than the scores in the first trial. There was also a statistically significant difference (F=7.86, p<0.001) between the test scores in the first trial of the test and retest. In this case, learning or carrying-over effects seemed to exert strong influence on the test–retest reliability coefficient. The presence of a learning effect on clients at level 5 during test–retest condition is not unusual (Allen and Blue, 1998). Learning effects may also occur with the short test–retest interval. Portney and Watkins (1993) suggested that the test–retest interval should be long enough to avoid learning effects but close enough to prevent changes in trait. The mean of the test–retest interval in this study was 28.3 days and the learning effect still persisted. To sum up, the comparatively low value of the test–retest reliability coefficient of the Cantonese version of the ACLS seemed to be primarily a result of testing effects.

Validity of a newly developed instrument raises huge concern from researchers and clinicians. The applicability of the ACLS to another cultural group is still unknown. In this study, no participants had reported having previous experience in performing leather lacing but some commented that the test’s tasks were similar to traditional Chinese needlework. Chinese needlework is a common task for the female Chinese population. In addition, task-oriented groups involving needlework are not uncommon in occupational therapy departments of psychiatric hospitals or psychiatric day hospitals in Hong Kong. Velligan et al. (1995) stated that the ACL – 90 relies on non-verbal problem-solving abilities and may be less culturally biased than verbal tests. Further investigation of the validity of the test is highly recommended.

There were several limitations in this study. The homogeneous sample from a single setting and the reliability coefficients generated by a set of scores with narrow range reduce the generalization of the results. Further research should examine the psychometric properties and clinical application of the Cantonese version of the ACLS.
Conclusion

The ACLS is built on the Cognitive Disability Model. Since its extensive use in different countries, the test has been translated into several languages by clinicians or researchers. The psychometric properties of the translated versions of the test were explored before applying the test clinically. Evidence gathered in this study showed that the Cantonese version of the ACLS seemed to be reliable despite methodological limitations in the study, such as small sample size and homogeneous sample. The results indicated that the test has the potential to be a reliable and valid clinical test for occupational therapists in Cantonese-speaking countries. Future research should be directed to examine the psychometric properties of the test and its clinical use in occupational therapy.

Acknowledgements

This study was conducted in partial fulfilment of the requirement for the degree of Master of Occupational Therapy from the Tufts University by Christopher Chan. I thank Diana Bailey, EdD, for her critical comments, thoughtful discussion of concepts and assistance in this research. I also thank all of the participants and the occupational therapists who participated in this study, especially Deborah Wong, MBA, OTR/L. Appreciation is extended to Durwood Marshall for his input on statistical analysis.

References


Address correspondence to Christopher Chan, Take Your Way, G/F, Bean Goose House, Sha Kok Estate, Shatin, Hong Kong.