VALIDATION OF
MILLER ASSESSMENT FOR PRESCHOOLERS
(CANTONESE VERSION)

BY

MAGDALENE YAN CHE POON

A THESIS SUBMITTED TO THE

RESEARCH AND POSTGRADUATE STUDIES OFFICE

IN PARTIAL FULFILMENT FOR THE

DEGREE OF

MASTER OF PHILOSOPHY

DEPARTMENT OF REHABILITATION SCIENCES

THE HONG KONG POLYTECHNIC UNIVERSITY

HONG KONG SAR, CHINA

MARCH, 1999

STATEMENT OF SOURCES

The idea of the present investigation and planning of the experiments were resulted from discussion between the author and Dr. Chetwyn Chan and Prof. Christina Hui-Chan.

All experiments in the present investigations were completed solely by the author.

The author declares that the work presented in this thesis is, to the best of the author's knowledge and belief, original, except as acknowledged in the text, and that the material has not been submitted, either in whole or in part, for a degree at this or any other University.

Magdalene Yan Che, POON
March 1999

DEDICATION

I dedicate this dissertation to God, for his love and support in my life; and my husband, Pak-cheong Ho and my three children, chun-yin, cheuk-yin and charm-yin. Their smiling faces have given me great support throughout the entire course of my master study.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank people who had been very supportive and helpful throughout the study.

The most important of all is my supervisor and co-supervisor, Dr. Chetwyn Chan and Prof. Christina Hui-Chan. They have given me expert guidance and opinions throughout the entire research. Their supervision has been intensive and superb.

I have to thank the expert panel members: Lenzs Lin, Lydia Leung, Winnie Chan, Sally Liu, Anita Fong, Miranda Ng, Joshua Mak, May Cheung, Ivy Wong, Irene Kwok, Betty Luk, Cynthia Woo who have spent their valuable time to share their precious clinical experience in working with children.

I would also like to thank Professor P. Fletcher and Dr. C. S. Leung of the Speech and Hearing Pathology Department, University of Hong Kong who have given me valuable assistance in designing the language specific items.

Lastly, sincere thanks are sent to all the children and the kindergartens who have volunteered to participate in this study. Without their support, this study will not be able to perform.

Abstract of thesis entitled "Validation of Miller Assessment for Preschoolers (Cantonese Version)" submitted by Magdalene Yan Che Poon for the degree of Master of Philosophy at The Hong Kong Polytechnic University in March, 1999.

Abstract

Miller Assessment for Preschoolers (MAP) is one of the popular assessment instruments used in Hong Kong for identifying children from two years nine months to five years eight months for developmental delays. This study aimed at translating three of its sub-scales: foundations, coordination, and verbal into Cantonese version (CMAP) in three age groups (II, IV and VI) and establishing the psychometric properties of CMAP. The translation and validation process involved direct and backward translation; test of equivalence between the two versions; design of new items; collection of content-related evidence by expert panel reviews; estimation of inter-rater reliability; and evaluation of construct and structural validity from a major field test. The panels consisted of six occupational therapists with an average 8.3 years of experience and six speech therapists with an average 5.2 years of experience. A total of 120 Chinese children with equal proportion of male and female in the age range from three years three months to five years eight months participated in the preliminary field test. In addition, a total of 184 children were selected by cluster random sampling from ten districts of Hong Kong to participate in the major field test. Children were 91 male and 94 female with mean age of 3.4, 4.5 and 5.5 years in their respective age groups of II, IV and VI.

Result from the expert panel reviews demonstrated that the items in CMAP were relevant and representative for identifying developmental delay

among Chinese children in Hong Kong. However, some of the panel members commented that the hand-to-nose and Romberg items were too easy for the children. High inter-rater reliability was estimated on the 18 items in CMAP (r = 0.80 to 1.00). Item difficulty and discriminative indices indicated that items: hand-to-nose, kneel-stand, Romberg, and stereognosis were too easy for Hong Kong children. When compared with their United States counterpart, findings revealed that children in Hong Kong tended to commit less error in the walks line test which can be explained by the difference in child rearing, parental expectation and age of schooling between Hong Kong and United States. Construct validity of the CMAP reflected that the verbal sub-scale was unidimensional and homogeneous. However, the foundations and coordination sub-scales had a less clear test structure which appeared among two to three different factors. In conclusion, the CMAP is recommended to be used as a screening instrument for children in Hong Kong. However, further studies are suggested to adjust difficulty levels of some of the items. The structural validity of some of the sub-scales require further empirical testing. Predictive validity and sensitivity of CMAP are feasible studies to be conducted to enhance the best utilization of the assessment instrument for children in Hong Kong.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICE

CHAPTER I
INTRODUCTION

This chapter begins with a statement of purpose that summarizes the objectives of this research study. It is followed by a brief description on the background and justification of the study, and finally, a summary in the content of this dissertation.

Statement of Purpose

This study proposed to translate three of the sub-scales in the Miller Assessment for Preschoolers (MAP) from English to Cantonese, and established the psychometric properties for the Cantonese version (CMAP). The objectives of the study were:

1. translate the instructions of foundations, coordination and verbal sub-scales of the Miller Assessment for Preschoolers into Cantonese (CMAP);
2. translate the test content of four language specific items: articulation, follow direction, general information, and sentence repetition into Cantonese;
3. collect evidence on the content validity (equivalence, relevance and representativeness) of the three sub-scales of CMAP by expert panel review;
4. evaluate the structural and construct validity of the CMAP by field testing 184 subjects in three age groups:
    - Age Group II: 3 years 3 months to 3 years 8 months
    - Age Group IV: 4 years 3 months to 4 years 8 months
    - Age Group VI: 5 years 3 months to 5 years 8 months

Background and Justification of the Study

Miller Assessment for Preschoolers (MAP) is an instrument developed by an occupational therapist aiming at identifying preschool children with mild to moderate developmental delay. Developmental delay among children is the failure in acquiring an ability at the expected time (Coker, 1989). Early screening of preschool children for developmental delay can predict risks for later developmental problems and determine the child's individualized training program as early as possible (Capute & Accardo, 1996; Fallen, 1985). The MAP is one of the commonly used assessment instruments for the preschoolers (Aylward, 1994; Banus, 1983; Daniel & Bressler, 1990; DeGangi, 1983; Deloria, 1985; Humphry & King-Thomas, 1993; Lane, Attanasio, & Huselid, 1994; Michaels, 1985).

The MAP was developed in 1980 by extensive expert panel reviews and item analysis to select the 27 items among 400 items in the initial pool. Its normative data was collected by stratified random sampling to 1200 subjects from nine geographic regions of the United States. The inter-rater reliability and test-retest reliability were high. The sensitivity of the MAP to detect children with developmental delay range from 0.54 to 0.61 to the WISC-R and the Woodcook-Johnson Psychoeducational Battery (Humphry & King-Thomas, 1993). Lemerand (1985) also in a one year follow up study found that the MAP had high sensitivity to teacher rating (0.83), referral to support services (0.70), and special placement and retention in class (0.70).

In Hong Kong, the MAP is commonly used among occupational therapists working with children (Child and Adolescent Psychiatric Working Group, Occupational Therapist Coordinating Committee, 1994). It is unique in

providing information to occupational therapists in testing sensory motor, cognitive and language functions of the children (Aylward, 1994; Miller, 1988). Currently, there are limited standardized assessment instruments validated for the therapists' operational use in Hong Kong. Thus, translation and validation of CMAP is essential. A group of occupational therapists had translated two sub-scales (non-verbal and complex tasks) of the MAP into Cantonese (Child and Adolescent Psychiatric Working Group, Occupational Therapist Coordinating Committee, 1994). The validation and translation of the remaining three sub-scales (foundations, coordination, and verbal) thus completes the translation of CMAP.

Test translation involves adaptation of the instrument with reference to culture, content and wordings that is needed in addition to the language translation of the test (Geisinger, 1994). In translating the three sub-scales of MAP, it involved the translation of test content of the verbal sub-scale of the children. An extensive study to language structure of Cantonese was required in order to reconstruct the test content in the ability in follow instruction, general vocabulary understand, and sentence structure of the preschool children. Through this process, it is therefore hoped that basic translation principles in adopting developmental assessment instruments to children in Hong Kong involving language abilities can be postulated.

In addition, the cultural differences in child development has always been an interesting issue to professionals working with children. It can affect a child's perceptual, cognitive, personality and moral development. In validating the CMAP, it included collection of normative data to 184 children by cluster random sampling to ten districts in Hong Kong. Comparisons were made with

the percentile scores of the United States children to have a gross overview on their similarities and differences in performance.

Moreover, the psychometric properties of CMAP in the inter-rater reliability; item difficulty and discrimination of the items in the three sub-scales with three age groups; and the construct of the three sub-scales were evaluated. Normative data was collected to children in Hong Kong. It is hoped that the CMAP can be applied to identify Hong Kong children with developmental delays so that early intervention to these children can be ensured.

Organization of the Chapters

This dissertation consists of six chapters. Chapter I is the introduction chapter. Chapter II is the literature review in which children with developmental delay; the assessment principles; and the importance of translating and validating standardized instruments will be discussed. Chapter III describes the method and procedures involved in data collection process of the validation of CMAP. Chapter IV reports the results of the test of equivalence; expert panel reviews; item difficulty in language specific items; and construct validity of CMAP. Chapter V discusses the interpretation of these findings. Chapter VI concludes and suggests implications of the findings to the application of CMAP in clinical practice.

CHAPTER II

LITERATURE REVIEW

Introduction

This chapter describes the literature review on the validation of the Miller Assessment for Preschoolers (MAP), the theoretical framework of the MAP and the validation theory. It begins with the definition and concept of children with developmental delay, its relationship with occupational therapy, and the assessment and intervention of these children. Next is a description of the MAP which aims at assessing children with developmental delay, the theoretical framework of the test construct, and the psychometric properties. Lastly, the theory and method of translating a standardized instrument in other countries is reviewed.

Children with Developmental Delay

Developmental delay is "the failure to acquire an ability at the expected time" (Coker, 1989, p.1).The most common presentation of developmental delay was the failure to achieve age appropriate developmental skills (Meisels, 1989; Shapiro, Palmer, & Capute, 1987).The prevalence of developmental delayed children in United States was about 10% (Rubin & Crocker, 1989). The common diagnosis of developmental delayed preschool children includes speech and language impairment, specific learning disabilities, and mental retardation. Children born with physical handicaps such as cerebral palsy, Down's Syndrome and blindness are easily detected by paediatricians before one year old. However, as children get older, skills performed became more

complex. Thus, the problems of developmental delay such as mild to moderate mental retardation, speech and language delay, and specific learning disabilities became more prominent when the child starts to learn at play groups, nurseries and kindergartens.

In Hong Kong, 3% of children (0 - 15 years old) have a disability according to statistic reports from the child assessment centres, the major centres for the provision of child developmental assessment in Hong Kong. Similar to the statistics in the United States, the common diagnosis among children with developmental delay included speech and hearing handicaps, specific learning difficulties and mental retardation. Motor clumsiness or motor delay, and autism were also among the most common diagnosis made by the child assessment centres (Lam, Yau, & Mak, 1996).

Although problems such as learning disabilities and mental retardation have no cure, early intervention in teaching the child and parents methods of compensating for the disability is emphasized (Batshaw & Perret, 1992). With the increase in awareness of the importance of early intervention, the provision of special education programs for children has expanded, in Hong Kong, from school age to preschool years and new born babies (Lam, Yau, & Mak, 1996). As a result, the Department of Health of the Hong Kong SAR Government has provided screening tests, free of charge, for normal children from new born to three years old. Those who failed this screening test were referred to child assessment centres for detailed assessment. Treatment and special preschool or school placement is then provided to those in need. These services includes early education training centres, integrated nurseries, special child care centres and specialty clinics run by the Hospital Authority in

Hong Kong (Opper, 1993). In 1989, there were 102 integrated institutes for preschool children and 14 special child care centres (Opper, 1993). In 1995, the child assessment centres received over 2,000 new referrals (Lam, Yau, & Mak, 1996).

Occupational Therapy and Children with Developmental Delay

Occupational therapists have a long history of assisting people with disabilities to achieve maximal independence in their occupations (Hall, 1918; Slagle, 1921). For children, occupational therapists assess the child's current developmental status and functional performance in the physical, sensorimotor, perceptual, self-help, and adaptive behavior domains (Allen & Pratt, 1989; Ottenbacher, 1991). Occupational therapists play a major role in the multidisciplinary team in assessment and early intervention of children with developmental delay (Allen & Pratt, 1989; American Occupational Therapy Association 1973; Ayres, 1963; Clark, Mailloux, & Parham 1985; Education for Handicapped Law Report:1986-1987). By analyzing the specific delays in sensorimotor, cognitive, language, social and self care aspects, therapists design treatment programs to promote the acquisition of age-appropriate skills for these children (Mailloux, Knox, Burke, & Clark, 1985). Thus, the children's maximal independence is enhanced.

Assessment of Children with Developmental Delay

Assessment was defined as making an evaluation or estimation of development which facilitated a clinical decision as to what intervention would be appropriate. It determined the existence of a delay or disability (Aylward,

1994). By delineating the child's abilities and comparing to those of the "normal" child, the examiner was able to develop a profile of strengths and weaknesses in the developmentally delayed children (Shapiro, 1996). This profile could assist the therapists to formulate the treatment plan to train the child's weak areas. Assessment could also screen developmental disorders or predict risks for later developmental problems (Capute & Accardo, 1996). In addition, assessment could determine the child's eligibility for special education service, and the nature required, and/or provide information required for individualized training programs (Fallen, 1985).

Framework in Identifying Developmental Delay

The assessment for a developmentally delayed child is done by comparing the performance of the child with reference to the normal developmental sequence. This was based on the belief that this sequence is closely linked to the maturation of the central nervous system (Shapiro, 1996). In addition, it is believed that development follows a specific sequence. Thus, by comparing the rates of attaining developmental milestones to those of norms it provides a means of early identification of developmental delay (Batshaw & Perret, 1992).

Areas of Assessment

In the early decades of the 20th Century, Gesell identified various areas of child development by relating the milestones to chronological age (Gesell & Ilg, 1943). He derived the main streams of development as gross motor, fine motor, visuomotor problem solving (or perceptual), expressive language, receptive language, and social and adaptive skills (Capute & Accardo, 1996). Basing the construction of assessments into these areas enabled therapists to

interpret the result in terms of profiles of strengths and weaknesses of children.

Gesell's definition of child development has been largely maintained by the practitioners and therapists in recent years. Aylward (1994), Brown and Elksnin (1994), Capute and Accardo (1996) and Shapiro (1996) also follow Gesell's principles of child development. As they designed the assessment areas, they also included motor (gross and fine) and language aspects. On the other hand, the visuomotor problem solving ability was replaced by a more global term: cognitive ability and the social and adaptive skills were described as social skills. These assessment areas will be discussed in detail in the following paragraphs. In addition, specific problems in children with mental retardation and specific learning disabilities, the two common diagnosis in children with developmental delay, will be illustrated.

Motor (Gross and Fine) Development

Motor development refers to the way in which children use their bodies. It is categorized as gross motor (involving the large muscles of the body) and fine motor (involving the small muscles of the body). Gross motor relates to locomotion, while fine motor activities relates to manipulation and eye-hand coordination (Turner & Hamner, 1994). The motor functions can only work smoothly with good sensory processes. The sensory components include visual, auditory, tactile, olfactory, and taste. These components are the primary components for learning (Turner & Hamner, 1994). As a child grows, the brain begins to integrate these sensory input with motor development. Thus, the first seven years of life are called the years of sensory-motor development (Ayres, 1979). To be specific, the visual and tactile senses

have a major effect in motor development. Also, sensory integration which also facilitates cognitive, language and social areas in a different dimension will be discussed separately.

With respect to the two most common diagnosis in children with developmental delay, children with mentally retardation generally lagged behind their normal peers in motor skills. On the other hand, learning disabled children's motor dysfunction usually related in deficits in directionality, laterality, motor coordination, balance, rhythm, body image, spatial awareness, and motor planning (Kalakian & Fichstaedt, 1994).

Cognitive Development

The foundation of cognitive development was developed by Jean Piaget. The preschool child (age three to seven) reached the preoperational stage, according to Piaget's Cognitive Development theory, with active use of language and symbols, trying to classify and group objects, though were not yet proficient (Piaget, 1926). Cognition referred to knowing and understanding. Piaget (1926, 1936 & 1970) believed that infants developed cognitively by exploring their environment with their bodies. They selected from an array of sensory input around them, and once the sensory input was received, they processed it to give it meaning and coherence. The cognitive skills for preschool children includes visual and auditory perception. Perception was the cognitive process that occurs when sensations are transformed into information (Turner & Hamner, 1994). The visual and tactile perception referred to the visual memory, figure ground, stereognosis and spatial relationship (Umansky, 1985). In addition, the child also increases their capacity of short term memory (Umansky, 1985).

In the case of the children with mental retardation, they have problems in memory and attention which affects their acquisition of information and the ability to relate the stimulus to other sensory input. They also perform only fair in storage and retrieval of auditory and visual information (Umansky, 1985). On the contrary, learning disabled children have variations in their strengths and weaknesses but some may have difficulties in the use of language, concept formation and memory (Rourke & Do Ho, 1994).

Language Development

Language is a uniquely human characteristic. There is no complete agreement on how children learn to talk (Turner & Hamner, 1994). It was generally agreed that both the biological basis and the environmental effects accounted for language development (Turner & Hamner, 1994). Language abilities include performance in phonology (articulation), morphology (minimal meaningful units of speech), syntax (grammar) and semantics (meanings) and pragmatics (psychosocial dynamics around which the use of language is based) (Easterbrooks, 1985).

Not only do children with language disorders have deficits in these areas, children with mental retardation also show a slower rate in learning all these concepts (Easterbrooks, 1985). Also, learning disabled preschoolers often experience word retrieval problems (Wiig & Semel, 1980) and exhibit poor listening skills (Easterbrooks, 1985).

Social Development

Social development is governed by a genetic structure that is common to all people but depends upon interaction with significant individuals and the environment. These interactions varied from culture to culture (Erikson, 1963).

Whereas motor milestones are almost entirely neuromaturationally determined and are influenced only minimally by the environment, social behaviors and self-help (adaptive) skills depend heavily on environmental factors such as social expectations, level of parenting skills, education, and training (Capute & Accardo, 1996). The performance of a child is culturally diverse and depends only minimally on biologic factors (Wagner & Stevenson, 1985). Social development relates to cognitive, motor and language skills but such milestones should never provide the sole support for a developmental diagnosis. It can only be used to confirm the overall pattern yielded by the rest of the developmental assessment (Capute & Accardo, 1996).

While assessment instruments designed for developmental delay mainly assess the motor, cognitive, language and social aspects, individual standardized assessment instruments differed in terms of the objectives of the assessment, the target subjects and the discipline using the instrument.

Common Developmental Assessment Instrument

Some of the commonly used assessment instruments used included the Denver Developmental screening Test, Bayley scales of Infant Development, Gesell Developmental Schedules, Kent Infant Development Scale and Miller Assessment for Preschoolers (Aylward, 1994). The age range covered in these assessment instruments ranged from 0 to 30 months and 0 to 6 years. The assessment areas were similar: cognitive, motor, and language aspects. Some also included the assessment of behaviour and sensorimotor areas. Among them, the MAP was unique in its age range and type of assessment (being a detail assessment rather than a screening test) (Aylward, 1994).

In local practice, the common assessment instruments used for the preschoolers included Griffiths Mental Development Scale, Visual Motor Integration, Reynell Language Developmental Scale, Wechsler Preschool and Primary Scale of Intelligence-Revised. In assessing the motor aspect of the children, therapists working in child assessment centres would use Bruininks-Osteresky Test of Motor Proficiency, Peabody Motor Developmental Rating Scale, and Movement ABC. Review of the literature indicated that the Reynell Language Developmental Scale was the only Chinese assessment instrument which could be applied by therapists underwent validation for the local use. Thus, the validation of the Cantonese Miller Assessment for Preschoolers and comparing the plaucible cultural differences would provide very much benefit the assessment and interrelation children with dysfunction in motor, cognitive and language performances in Hong Kong.

Miller Assessment for Preschoolers (MAP)

The MAP was developed by Lucy Jane Miller, an occupational therapist, in 1982. The purpose of the assessment was to identify children with mild to moderate developmental delay (Miller, 1988). The test was designed for preschoolers with age range from 2 years 9 months to 5 years 8 months divided into six age groups. The age group and age distribution are shown in Table 2.1.

Table 2.1

Age Distribution of the Age Groups in MAP

Age Groups	Age Range
I	2 years 9 months to 3 years 2 months
II	3 years 3 months to 3 years 8 months
III	3 years 9 months to 4 years 2 months
IV	4 years 3 months to 4 years 8 months
V	4 years 9 months to 5 years 2 months
VI	5 years 3 months to 5 years 8 months

In addition, the MAP also resolved the children's presenting problems with a profile made up of five sub-scales. This profile assisted therapists in differentiating children's strength and weakness. Based on this, a tailor made treatment programme can be made for the child (Miller, 1988). The five sub-scales in MAP included: foundations, coordination, verbal, non-verbal and complex tasks. The foundations sub-scale mainly tested the child's sensory motor functioning. The coordination sub-scale tested the child's eye hand

coordination and oral motor coordination. The verbal sub-scale tested the language ability of the child. The non-verbal sub-scale tested the child's cognitive functioning. The complex tasks sub-scale mainly focused on testing the child's integration of cognitive, sensory and motor functions. The test was constituted by 27 items grouped under these five sub-scales. The detail description of each item is shown in Appendix A.

The initial item pool was developed by a comprehensive literature review. Some items were adopted and modified for the preschool age range from existing tests including the Southern California Sensory Integration Test (Ayres, 1972), Meeting Street School Screening Test (Hainsworth & Siqueland, 1969), Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy, & Kirk, 1968), and Bruininks-Oseretsky Test of Motor Proficiency (Bruininks, 1978). Other items which were selected from the Gesell Developmental Schedules (Gesell, Halverson, Ilg, Castner, Ames, & Amatruda, 1940) , Detroit Tests of Learning Aptitude-2 (Hammill, 1985), and the Columbia Mental Maturity Scale (Burgemeister, Blum, & Lorge, 1957), were revised by Miller so as to be more contemporary. These preliminary draft of items was discussed among paediatric experts including paediatricians, occupational therapists, physiotherapists, speech therapists, psychologists, and special education teachers (Miller, 1988). The original pool of about 400 items was administered to 400 children in Massachusetts public schools' preschool screening program. Items which did not meet the test construction standards in terms of the item difficulty, item discrimination and correlation studies were deleted. Item difficulty, item discrimination and correlation studies referred to the percentage of children passing each item at each age, the ability of each item to

differentiate between normal and at-risk groups of children and point-biserial correlation of each item with sub-scales and total score. Meanwhile, new items were devised to replace the deleted items. This exercise was repeated twice on children in the Tewksbury, Massachusetts public schools; children in Boston who were diagnosed to have developmental delay; and children in Walpole, Massachustts schools. The total number of subjects involved was 1,014. The revised pool of items was then administered to about 500 randomly selected children in nine geographical regions. The socioeconomic status of the child's family, educational and professional classification of the parents and income of the family closely resemble the United State Census. Data analysis of these items was done. The 27 items were selected based on six criteria:

- ability of the item to discriminate between age groups,
- ability of the item to discriminate between normal children and children with developmental delay,
- representativeness among the sub-scales,
- simplicity of the equipment needed,
- ease of administration, and
- content validity comments from the expert panel.

The MAP was standardized in 1980 by stratified random sampling to select 1200 subjects from nine continental geographic regions of the United States. The inter-rater reliability of the MAP was tested on 40 randomly selected children. The correlation coefficient of the raters in rating the five sub-scales ranged from 0.84 to 0.99. However, the inter-rater reliability of individual items in the sub-scales were not published. Moreover, a relatively

low correlation coefficient was found in the coordination sub-scale. In order to evaluate the relatively low correlation coefficient, Milller (1988) listed the percentages of agreement in the items of coordination sub-scale. All items' percentages of agreement among the raters were above 90% except the articulation item in which the percentage of agreement was 77.2%. Thus, the relatively low correlation coefficient was mainly due to the low consistency of the raters in the articulation item since high consistency was found among the other items in the coordination sub-scale.

The test-retest reliability studies were conducted on 81 normal children. The test retest stability of the score in the sub-scales were between 72% to 94% (Miller, 1988). The coordination sub-scale was the lowest while the other sub-scales were over 80%.

The content validity of the MAP was reviewed by an expert panel consisting of paediatricians, psychologists, occupational therapists, physiotherapists, speech therapists, and special education teachers. The construct validity of the MAP was reviewed showed by varimax rotated factor matrix to the 27 items of the MAP. Results reviewed that the factor loading of the items in each factor was low except three items under verbal sub-scale (r = 0.61 - 0.69). The foundations sub-scale did not cluster together, nor did they cluster with the items from other sub-scales. The coordination sub-scale loaded on two factors with low factor loading (r = 0.33 to 0.59). There were items which did not load distinctly on any factor (tower, finger localization, object memory, vertical writing, Romberg, articulation and digit repetition).The correlation of items with the total score ranged from 0.20 to 0.52 (p<0.01). The

correlation of the MAP total score with the five sub-scales ranged from 0.65 to 0.78 (p<0.01) (Miller, 1988).

The predictive validity involved 338 preschoolers who were initially tested in 1979-1980 and subsequently evaluated four years later. Moderate correlation was found between the MAP total score and WISC-R scales with correlation coefficient: 0.45 to 0.50 (p<0.001). However, low correlation was found between the MAP total score and Woodcock-Johnson Math, Reading, and Language measures with the correlation coefficient 0.35 to 0.38 (p<0.001) and Bruininks-Osterek Test of Motor Proficiency, correlation coefficient: 0.39 (p<0.001) (Miller, 1988). Schouten and Kirkpatrick (1993) also doubted the predictive validity of the MAP. On the other hand, Humphry and King-Thomas (1993) had a different opinion. They commented that poor correlation was caused by high drop out rate (338 out of 800 subjects) especially those coming from the lower socioeconomic class. This made good predictive validity results impossible. On the other hand, the sensitivity of the MAP to detect children with developmental delay range from 0.54 to 0.61 to the WISC-R and the Woodcook-Johnson Psychoeducational Battery (Humphry & King-Thomas, 1993). Lemerand (1985) also in a one year follow up study found that the MAP had high sensitivity to teacher rating (0.83), referral to support services (0.70). special placement and retention in class (0.70).

The MAP was widely used among occupational therapists all over the world. In Australia, Reid (1987) conducted a survey and the MAP was used in 60% of the occupational therapists working with children in Australia. Translation of the MAP in other countries such as Canada, Denmark, Finland, Israel and Japan were done. It was recognized as one of the best available

assessment instruments for developmental delayed children (Banus, 1983; Campbell, 1989; Daniel & Bressler, 1990; DeGangi, 1983; Deloria, 1985; Humphry & King-Thomas, 1993; King-Thomas & Hacker, 1987; Lane, Attanasio, & Huselid, 1994; Michaels, 1985).

Theoretical Framework underpinning MAP

The theoretical foundation of the MAP was based on literature from child development, education, psychology, language development, physiotherapy, occupational therapy, and medicine (Miller, 1988). The MAP covered the areas of sensory and motor integration, cognitive, language, and the integration of all these aspects. The assessment areas of the MAP basically matched the general assessment areas of Gesell (motor, cognitive, language and social) of the developmental delayed children. However, the five sub-scales under the MAP did not directly match these four areas. The foundations and coordination sub-scales tested the motor function of the child. The non-verbal and verbal sub-scales tested the cognitive and language functions. The complex tasks sub-scale tested the integration of cognitive and motor function. The social aspect was not included in the assessment. On the other hand, a behavioural observation checklist was included for the therapists to evaluate the overall social behaviour of the children. The theoretical framework of three sub-scales: foundations, coordination and verbal which were selected in this research is discussed in the following paragraphs.

The foundations sub-scale mainly assessed the sensory motor function of the child. Sensory motor development referred to the qualitative changes when children learn and master early capabilities by interaction between themselves and the environment (Dunst, 1998). Under the theoretical

framework of MAP, the sensory motor function included the child's sense of position and movement function (the integration of the motor and the vestibular, proprioceptive, and cerebellar senses); tactile function and the development of movement patterns (flexion, extension, weight shifting and rotation). The position and movement function included the measurement of the ability to maintain the position with the eyes closed or while stepping and writing with vision occluded (items: Romberg, stepping, vertical writing); the estimation of the joint movement with vision occluded (item: hand-to-nose); the balance mechanism (items: walks line and kneel stand). The tactile function included the finger differentiation with vision occluded (item: finger localization); and the recognition of forms and objects by touch (item: stereognosis). The development of movement patterns included the integrative movement of the whole body in balance (item: supine flexion) and stepping (item: rapid alternate movement) (Miller, 1988). Some of these items involved the test of neuromotor function of the children which was usually included in part of the motor assessment. The diadochokineasia (ability to perform rapid alternate movement), graphesthesia (ability to recognize object by feeling with hand), finger localization, and toe tapping (ability to perform rapid alternate movement with feet) (Montgomery, 1996). Miller (1988) found that the items under the foundations sub-scale were not measuring a single domain of behaviour as they did not cluster together. Montgomery (1996) also shared this opinion although he also placed these sensory motor tests under the motor aspect of the child's development.

The coordination sub-scale involved the gross motor coordination in locomotion (item: walks line, rapid alternate movement); fine motor

coordination (item: tower, motor accuracy, vertical writing) and oral motor coordination (item: tongue movement and articulation) (Miller, 1988). There were no distinct differences in some of the items compared to the foundations and coordination sub-scales. For example, Miller (1988) suggested that walks line, rapid alternate movement and vertical writing items were overlapping in both sub-scales. Not all the items under coordination sub-scale were clustered together in factor analysis. The items: kneel stand, rapid alternate movement, tongue movement clustered together (r = 0.37 - 0.43). Other items were clustered under other factors randomly. The correlation of items: tower, articulation and vertical writing were very low ($r < 0.30$) (Miller, 1988). Current literature mentions about the importance of the articulation test to assess the phonological function of the children (Ng, 1995). However, the tongue movement, kneel stand, and vertical writing test in the motor aspect of child developmental assessment are not mentioned in recent literature.

The verbal sub-scale assessed the language ability of classification, association, general knowledge, follow directions, understanding quantitative and prepositional concepts and sentence imitation ability (test items: general information, follow direction, sentence repetition and digit repetition) (Miller, 1988). Montgomery (1996) also suggests to use sentence repetition and digit repetition to assess the auditory memory of the children. Furthermore, he suggests comprehension and information testing which is very similar to the general information and follow direction items assessing the verbal ability of the children (Montgomery, 1996).

The detail breakdown of the MAP's items in relation to the areas of assessment of developmental delayed children is listed in Table 2.2.

Table 2.2

Assessment Areas in MAP

Items	Assessment Areas	Specific Areas
Tower	Motor	fine motor
Sequencing	Cognitive	problem solving
Block designs	Cognitive	spatial relationship
Block tapping	Cognitive	memory
Stereognosis	Cognitive	tactile discrimination
Finger localization	Cognitive	tactile discrimination
Object memory	Cognitive	memory
Puzzles	Cognitive	visual perception
Figure-ground	Cognitive	figure ground perception
Draw-a-person	Cognitive & motor	body scheme, coordination
Motor accuracy	Motor	fine motor
Vertical writing	Motor & cognitive	fine motor, position and movement
Hand-to-nose	Cognitive	position and movement
Romberg	Cognitive & motor	position and movement
Stepping	Cognitive & motor	position and movement
Walks line	Cognitive & motor	balance, gross motor
Supine flexion	Motor	motor planning
Kneel-stand	Motor	motor planning, balance
Imitation of postures	Motor & cognitive	visual spatial relationship
Tongue movements	Motor	oral-motor

(To be continued)

Table 2.2 (Continue)

Assessment Areas in MAP

Items	Assessment Areas	Specific Areas
Rapid alternate movement	Motor & cognitive	position and movement, gross motor
Maze	Cognitive	problem solving
General information	Language	language understanding
Follow direction	Language	language understanding
Articulation	Motor	Oral-motor
Sentence repetition	Language	language understanding and expression
Digit repetition	Language	memory

* shaded areas are items included in this research

Application of MAP in Hong Kong

Occupational therapists in Hong Kong working with children often use MAP as one of their assessment tools (Child and Adolescent Psychiatric Working Group, Occupational Therapist Coordinating Committee, 1994). The instrument was unique in assessing developmental performance of children at preschool age, it not only assesses the sensory motor aspect but also cognitive and language functions of the children.

However, since there is no Cantonese translation of the original MAP, therapists had to use their own interpretation of the English instructions. Thus, a group of occupational therapists had translated two of its sub-scales: non-verbal and complex tasks. The validation process further involved tests of equivalence and content validity of the two translated sub-scales and

assessed their inter-rater reliability, test-retest reliability and gender differences (Child and Adolescent Psychiatric Working Group, Occupational Therapy Coordinating Committee, 1994). Different from the non-verbal and complex tasks sub-scales, in another sub-scale of MAP: the verbal sub-scale, the test contents of articulation, general information, follow direction and sentence repetition items involved cultural related questions. For example, the baseball bat picture, a game rarely seen by Chinese preschoolers in Hong Kong, was found in general information item. And the English grammatical requirement of the sentence repetition item was not relevant to children in Hong Kong, predominantly Chinese. The differences in culture and language had a significant impact on the validity of the original English version (Geisinger, 1994; Satorius & Kuyken, 1994; Vernon, 1973). As a result, the translation and evaluation of the validity and reliability of the other three sub-scales of CMAP was essential before it could be used in assessing Chinese children in Hong Kong.

Culture could be viewed as a multifaceted influence which is learnt by direct and indirect daily experiences based on what people do (cultural behavior), said (speech messages), make and use (cultural artifacts). The child learns this life pattern of beliefs and values which shapes the way he or she believes, thinks, perceives, feels and behaves (Levine, 1987).

Child development in relation to a specific culture has been studied from different perspectives. The cultural difference could affect a child's development in perception, cognitive, personality and moral development. These cultural perspectives could be seen in the child's human and non-human environment. Human environment being the child's major caretaker

and parents while non-human environment includes the child's schooling and nutritional background (Wagner & Stevenson, 1985). The MAP which aimed at assessing children of preschool age in the sensory motor, coordination, cognitive and language functions by comparing the normal developmental sequence, were therefore unique across different cultures.

Validation was essential in translating the MAP for Hong Kong children. Hong Kong, is a new target population with a specific culture, and language background. In particular, both the human and non-human cultural background of Hong Kong children should be addressed.

Translation of Standardized Assessment Instruments

Test translation (or sometimes named test adaptation) involved adaptation of the test in references to culture, in content, and in wording that was needed in addition to the language translation of the test (Geisinger, 1994). Special attention should be paid to translating the standardized assessment instrument and applying it to different countries (Satorius & Kuyken, 1994; Vernon 1973). According to Satorius and Kuyken (1994), there are four approaches to translation: ethnocentric approach; pragmatic approach; etic plus emic approach and the interpretation approach. The first approach assumed that the conceptual dimension among the two cultures was transferable. Based on this, the source instrument (in this case, the MAP of United States) can be translated into the language of another culture without consideration of the appropriateness of the instrument in the target culture (in this case, the CMAP for Hong Kong). The second approach considered the cultural factor and only translated those parts of the construct found to be common to both cultures. The third approach looked for both the common

ground between cultures in the conceptual basis of an instrument (etic aspect) and the culture-specific facets of the concept in the target culture, with a view to operationalizing and also measuring them (emic aspects). The emic concepts, once identified, had to be assessed in separate studies in both cultures. The fourth approach involved the interpretation of items in the source instrument, for use in the target culture. If the distance between the construct of the two cultures was too large, the process may result in two similar instruments whose application would not yield comparable data (Satorius & Kuyken, 1994). In translating MAP for children in Hong Kong, the etic plus emic approach was suggested. This was because when comparing the culture of Hong Kong and the United States, there was some common ground as well as some cultural specific aspects.

Cultural Background - Hong Kong Versus United States

For the etic aspect, Hong Kong Chinese have always been recognized as "westernized" people (Bonavia & Sterry, 1994) after being part of the British Colony for over a hundred years (Rodwell, 1991). For the emic aspect, a traditional complicated Confucian ethic of social inter-relationships, which has dominated all Chinese societies since the 5th Century BC, still exists among the Chinese (Bonavia & Sterry, 1994). These similarities and differences will be discussed in terms of the three major culture facets: parenting, nutrition, and schooling.

Parental Influence

Child rearing is an important function of the Chinese family. The Chinese first nurture and protect the infant, then train the older children to bring honor to himself and the family by educational achievement, effort and good conduct

(Char, 1981). Parents in Hong Kong were conformed to the traditional Chinese pattern of socialization which emphasizes obedience, proper conduct and impulse control (Cheung, Lam, & Chau, 1990; Mann, 1984). "Perseverance", which meant rule compliance was very important for Japanese and Chinese parents when compared with United States (Tobin, Wu, & Davidson, 1989). More emphasis would be put on the intellectual development as compared to United States (Tobin, Wu, & Davidson, 1989). In the teaching learning process, parents encouraged the practice of memory, which was a traditional way of learning centuries ago. In the past, most subjects were taught by rote memory by a tutor employed by the clan (Rodwell, 1991). Today, traces of this old tradition can still be found in some Chinese families (Mann, 1984; Tse & Ngau, 1995). With repeated practice and reinforcement in following rules and instructions; memory; and intellectual play, children in Hong Kong may sometimes perform better in tests which involve the following of instructions. The MAP, being a test in which every item involved following the instructions, may elicit a better result for children in Hong Kong. Besides, in items which involved memory and intellectual components, the Hong Kong children may have had previous experience with similar types of games. Thus, their performance may be better in items such as digit repetition, sentence repetition, and object memory.

In Chinese tradition, girls are not encouraged to have as vigorous motor activities as boys which causes a difference in motor performance (Lam, 1993). In terms of gender difference, a study on the motor performance of preschoolers in Hong Kong revealed that girls motor performance was lower than boys (Lam, 1993). The MAP primarily designed as a screening

assessment, emphasized that the test could not differentiate the difference in quality of the performance in normal children. Thus, there was no separated norm for different gender. Thus, the gender difference of the CMAP had to be confirmed.

Nutritional Background

According to the State of the World's Children 1995 (Grant, 1995), the percentage share of household income and consumption, percentage of one-year-old fully immunized children, primary school children enrollment ratio, and life expectancy between Hong Kong and United States were similar. This reflected that Hong Kong, in terms of the income, nutrition, and health conditions, were comparable to those of United States. With this evidence, it was concluded that the physical development of children in Hong Kong and United States were similar due to the similar socioeconomic background.

Schooling in Hong Kong

Due to the introduction of a selection process based on entrance examinations for primary school places, parents began to enroll their preschool aged children in kindergartens to prepare them for the examinations. (Hong Kong Government, 1979). Regular schooling is encouraged by most parents for their children at an early age (Opper, 1993). The 1991 and 1996 Census of Hong Kong reviewed that over 95% of children at ages three to five were at school while 100% of children at age six were at school (Hong Kong Government, 1996). The majority of preschool children go to a kindergarten or child care center for the purpose of education, care and supervision. The curriculum content were mainly in four developmental areas of physical, intellectual, language and social-emotional development (Opper, 1993). A

study found that the academic curriculum for the preschoolers tended to put much more emphasis on word recognition, copying and number concepts which was very close to that in the primary school compared with the United States (Tse & Ngau, 1995). This practice of memory and academic work at preschool age increases the children's level of performance in table tasks in the test. Moreover, since Hong Kong's teacher and student ratio remains large, teachers often request children to be very quiet and help themselves during class (Chow, 1993). This reinforced the assumption made earlier that children in Hong Kong had better compliance to instructions during the test. Therefore, for items involved counting errors made or best result with no time limit, Hong Kong children might perform better. These items included tower (counting number of blocks built) and Romberg (ability to stand still within 15 seconds), etc.

Validity and Model of Validity

The validity of a test is concerned with what the test measured and how well it did so (Anastasi & Urbina, 1997). Fundamentally, all procedures for determining test validity are concerned with the relationships between performance on the test and taking into consideration other independently observable facts about behaviour characteristics under consideration (Anastasi & Urbina, 1997). There are two major trends currently used in the validation of an instrument: the theoretical framework of the instrument and the linkage between the theory and verification through empirical and experimental hypothesis testing (Anastasi, 1992, 1995). Thus, the value of the "construct", the theoretical entity (Anastasi & Urbina, 1997) was the

fundamental concept of validity (Anastasi & Urbina, 1997). Construct validity was the degree to which an assessment instrument measures targeted construct (Haynes, Richard, & Kubany, 1995). It subsumed all categories of validity (Messick, 1993). Validation of the instrument involves the content-description procedures - content validity testing; criterion-prediction procedures - concurrent and predictive validity testing; and construct-identification procedures - factor analysis, internal consistency and convergent and discriminant validation (Anastasi & Urbina, 1997; Haynes, Richard, & Kubany, 1995). Measures of the content, predictive, concurrent, criterion-related validity and factor structure provides evidence about the construct validity Item analysis, internal consistency indices, and the obtained factor structure also provides essential information about the degree to which an item taped the intended constructs and facets (Smith & McCarthy, 1995). During this research, the content validity and the construct validity in terms of the internal consistency of the items to the sub-scales and the factor loading of the items will be studied. Thus, detailed descriptions of these concepts are discussed below.

Content-description procedures - Content Validity

Content validity was the degree to which elements of an assessment instrument were relevant to and representative of the targeted construct for a particular assessment purpose (Haynes, Richard, & Kubany, 1995). "Relevance" refered to the appropriateness of its elements for the targeted construct and function of assessment (Messick, 1993). "Representativeness" refered to the degree to which its elements were proportional to the facets of the targeted construct (Nunnally & Bernstein, 1994). Content validation could

be done by multiple experts on every element of the assessment instrument, using evaluation scales, on applicable dimensions such as relevance, representativeness, specificity, and clarity (Haynes, Richard, & Kubany, 1995). Close examination of the proportional representation of items should be made to ensure that items were not overrepresent or underrepresent facets of a construct (Anastasi, 1988). In addition, the technical quality (e.g. grammer, wording) should be reviewed (Hambleton & Rogers, 1991). For this research, expert panels were invited to evaluated the content validity of CMAP on the equivalence, relevance and representativeness.

Construct-identification procedures - factor analysis

Factor analysis is a refined statistical technique for analyzing the interrelationships of behaviour data and it is particularly relevant to construct-validation procedures (Anastasi & Urbina, 1997). In the process of factor analysis, the number of variables or categories in terms of which each individual's performance which could be described was reduced to a relatively small number of factors. These factors could be utilized in describing the factorial composition of a test, together with the loading of each major factor and the correlation of the test with each factor (Anastasi & Urbina, 1997; Clark & Watson, 1995; Smith & McCarthy, 1995). Factor analysis involved first, subjecting the items to a principal components analysis (Comrey, 1988) and extracting the first few factors and secondly, examining the loading of the items. Items that loaded relatively strong on one factor and weak on other factors were excellent candidates for retention (Clark & Watson, 1995). Following the same procedures, the inter-relationship of the items in CMAP to their respective sub-scales will be examined by factor analysis in this research.

Construct-identification procedures - internal consistency and item distribution

The degree of homogeneity of a test has some relevance to its construct validity (Anastasi & Urbina, 1997). Internal consistency and inter-item correlation reflected the homogeneity of the test to some extent (Clark & Watson, 1995). Before conducting more complex structural analyses, the response distributions of the individual items in terms of skewness and unbalanced distributions should be eliminated since they were likely to correlate weakly in the pool (Clark & Watson, 1995). The next crucial stage was to conduct structural analyses is by examining the internal consistency, the coefficient alpha. Internal consistency referred to the overall degree to which the items that make up a scale were intercorrelated (Clark & Watson, 1995). It gave some indication to the homogeneity of the items to the test or the sub-scales of the test (Anastasi & Urbina, 1997). However, the average inter-item correlation was also a useful measure to the homogeneity of the items (Clark & Watson, 1995). In this research, the coefficient alpha and the inter-item correlation of the items to their related sub-scales would be evaluated to test the homogeneity of the CMAP.

Process of Validation - Cross-Cultural Normative Assessment

In translating the MAP into a Cantonese version, special consideration had to given to its application for a new target population in terms of cultural and language differences. Specifically to the assessment instrument, its development was built upon the context of contemporaneous theories about the targeted construct (Haynes, Richard, & Kubany, 1995). However, the domain of the construct would evolve over time and degrade the content validity. When adapting a standardized assessment instrument for new target

populations which differs appreciably from the original population in terms of culture or cultural background, country, and language, validation is crucial (Geisinger, 1994). Geisinger (1994) thus suggested the steps in translating a standardized assessment and adopting it to a new population. Firstly, when translating the assessment, the translator has to be fluent in both languages, knowledgeable about both cultures and familiar with the characteristics and content measured. Secondly, expert panels should be formed to examine the translation quality of the assessment. Thirdly, pilot tests should be done on a small sample of individuals comparable to the eventual target population to investigate the understandability of the instructions by the target population, the acceptability of the administration time and the technical wordings of the test. Fourthly, there should be a field test to a large sample representative of the eventual population to be assessed with the assessment instrument. The internal consistency, inter-rater reliability and validation in terms of the content validity, construct validity and criterion validity should also be performed.

This research largely followed the steps suggested by Geisinger (1994) as mentioned above. In addition, backward translation was added to ensure the quality of the translation. Also, an additional preliminary field test was added to the language specific items so as to ensure a better pool of questions designed for these items before the actual field test. The validation process selected included the test of content validity and construct validity by examining the factor loading, item analysis, and internal consistency. In addition, the gender difference was examined so as to ensure that the cultural difference did not have an effect on the construct of the CMAP.

CHAPTER III

METHODOLOGY

Introduction

The procedures involved in the validation process of Miller Assessment for Preschoolers (Cantonese Version) (CMAP) followed the method and techniques proposed by Geisinger (1994) and Sartorius and Kuyken (1994) in principle. They included translation of the test; evaluation of equivalence and content validity of the translated version; revision of the translated version to suit language and cultural relevance; pilot test the instrument and finally establishing the instrument's psychometric properties. This research was divided into six phases:

I. Translation and evaluation of equivalence to the instructions of all 18 items;

II. Translation and evaluation of equivalence of the test content of four language specific items;

III. Evaluation of content validity of CMAP;

IV. Preliminary field test of the four translated items in Phase II;

V. Pilot test of CMAP; and

VI. Field test on construct and structural validity of CMAP.

Before all the translation done, formal approval was obtained from the Psychological Corporation who owned the copy right of the MAP.

Phase I: Translation and Equivalence

In converting MAP into CMAP, the English instructions contained in the original manual were translated. The step by step procedures of test administration were not translated because the therapists who used it were largely bilingual. The test content of the four language specific items (general information, follow direction, sentence repetition and articulation) , which required both contextural and cultural considerations were handled separately in Phase II.

Translation of the instructions

One qualified translator translated all the test instructions into Cantonese. Then, a backward translation of these instructions into English was made by another independent qualified translator who had working experience in the rehabilitation field for over one year. The two different versions were then evaluated by the researcher in order to ensure the compatibility of the meanings between them (Satorius & Kuyken, 1994). Modifications were then made to discrepancies found in the translated version which showed distorted meanings of the original MAP. Both translators chosen were fluent in both languages and knowledgeable about both Western and Chinese cultures.

Evaluation of Equivalence

Expert panels were formed to assess the fluency, semantic and technical equivalence of the instructions of the 18 items (Foster & Cone, 1995). This was regarded as an effective technique to ensure the quality of the translation (Geisinger, 1994). A questionnaire was designed to facilitate the panel members' evaluation during the review which assisted in identifying the elements that required further refinement (Haynes, *et al.*, 1995).

Sample

The two expert panels included six occupational therapists (Panel A) and six speech therapists (Panel B). Panel A was requested to evaluate the 14 items under foundations and coordination sub-scales. In view of the wide range of subspecialties of occupational therapy service, only those with more than eight years experience working with children were selected. Potential panel members were contacted by the researcher through telephone to explain the nature of the expert panel and to invite their participation. After obtaining verbal agreements, formal invitation letters (Appendix B) were sent out to obtaining confirmation. Panel B were speech therapists with experience of two years or more. They were requested to evaluate the four language specific items as they were more familiar with the standard procedures in assessing language functions. Local, bilingual speech therapists with over two years working experience and having Cantonese as their mother language were contacted. The recruitment procedures were the same as that for Panel A.

Data Collection Procedures

Members in the two panels separately attended a three-hour review meeting. Members were requested to evaluate the translated instructions of CMAP following two standardized questionnaires (Appendice C and D). The operational definitions used in the questionnaire were explained by the researcher before hand. The panel members were then allowed to complete the questionnaire in one hour. After completing the questionnaire, the panel members were asked to discuss their ratings and comments. In the discussion,

they were facilitated to seek consensus of possible modifications to the translation.

Instrumentation

The questionnaires (Appendice C and D) used in the expert panel review consists of a total of 28 questions for panel A and 8 questions for panel B. Panel members were requested to assess the fluency and technical equivalence (grammar and syntax) of the translated version and the semantic equivalence (denotative and connotative sameness of the words) of each item in the translated version. "Fluency and technical equivalence" referred to the grammar and syntax of the translated version (Satorius & Kuyken, 1994). A typical question was "Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?" "Semantic equivalence" referred to the denotative and connotative sameness of the words of the translated version (Satorius & Kuyken, 1994). A typical question was "Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?" In each question, the panel members had to rate their opinion under a four point Likert scale ranging from strongly agree (4); agree (3); disagree (2); and strongly disagree (1). In addition, members were requested to give their written comments on the questions which they rated as disagree or strongly disagree.

Method of Data Analysis

The mode of the rating of each item in the questionnaires was calculated. The percentages of agreement of the mode rating among all panel members were computed. Percentage of agreement (Po) was a measure of how often individual raters agreed on the ratings which reflected consistency.

$$Po = \frac{\text{number of exact agreement}}{\text{number of possible agreements}}$$

(Portney & Watkins, 1993).

As the questions utilized a four point Likert scale, a mode score at or above 3 indicated that most of the panel members agreed to the translation of that item whereas a mode score at or below 2 indicated that most of the panel members disagreed with the translation. The consistency of the mode score was reflected by the percentage of agreement. The closer the agreement was to 100%, the better the consistency of the comments of the panel members were. In addition, all questions which were rated on or below 2 by any of the panel members were discussed among the panel members to reach a compromised decision on retention or modification of the translation.

Phase II: Translation and Evaluation of Equivalence of Four Language Specific Items

The objective of this phase was to translate the test content of the four items (articulation, follow direction, general information, and sentence repetition) into Cantonese and evaluate the equivalence between the translated and the original version. In addition, newly designed questions for these four items were designed. Since the original test content of the four

items was based on the English phonology and grammar, translation of the test content into Cantonese was essential for its use in a Chinese society.

In the original MAP, three questions were set for general information, follow direction and sentence repetition items separately. In order to select the best three questions for items of CMAP, both direct translation of the test content from the original MAP and three newly designed questions were made in the general information and follow direction items. Thus, a total of six questions were made for the preliminary field test to select the best three. The new questions of the follow direction and general information items were designed based on the principles and protocols adopted in the Reynell Developmental Language Scale (Chinese version) (Ko *et al.*, 1987). For the articulation item, in the original MAP, a total of 33 phonemes were tested. In translating this item, guidelines set out in the Cantonese Segmental Phonology Test (So, 1993) was followed. Both of the instruments mentioned here were validated previously and normative data was available. However, for the sentence repetition item, literature reviews showed no previous work had been done in this area locally. As a result, the translation principles and procedures made reference to three major sources:

- the general Cantonese grammar structure (Matthews & Yip, 1994);
- the Reynell Developmental Language Scale (Chinese version): the age range of acquiring Chinese grammar; and
- Cantonese sentence repetition test constructed by Ng (1995).

In order to ensure that the best three questions could be selected, three questions were made by direct translation from the MAP and six newly designed questions were devised by the researcher. The entire process of

translating and designing the questions for the language specific items was done under the guidance of Dr. C. S. Leung and Professor P. Fletcher of the Speech and Language Pathology Department, The University of Hong Kong.

Test for Equivalence

The same members of panel B were invited again to evaluate the equivalence between the translated and the original version of the follow direction, general information and sentence repetition items. The data collection procedures and instrumentation used were similar to that in phase I. The data analysis procedures were also similar to that of phase I. On the other hand, the content validity of these language specific items, both the questions by direct translation of the MAP and the newly designed ones were evaluated in phase III.

Phase III: Content Equivalence of CMAP

This phase was aimed at gathering evidence on the content-related validity - relevance and representativeness of the 18 items for the assessment of verbal, coordination and foundations functions of children.

Panel members in panel A were gathered again to review items in the foundations and coordination sub-scales. In order to allow a realistic evaluation, each of the panel members were asked first to administer the CMAP to a total of five children so that they became familiar with the testing procedures, administration and experienced the response from the children. At the same time, the same members of expert panel B in phase I were invited again to evaluate the content validity of the four language specific items.

Data Collection Procedures

All the panel members, from both panels, separately attended an one hour orientation program which introduced the purpose and content of the CMAP. It also included demonstrations on the administration of the CMAP. Discussions were held to clarify queries concerning the theoretical framework and test domain. The panel members were then asked to attend another meeting two months later to complete a questionnaire (Appendix E) for evaluating the relevance and representativeness of the test content. During the meeting, the researcher explained the purpose of the meeting; the operational definition of relevance and representativeness; and the cultural factors which may affect the content validity of CMAP; and the method of answering the questionnaire. The members were requested to determine whether the items in CMAP were relevant and representative to the assessment of developmental delayed children in Hong Kong.

Instrumentation

The questionnaire consisted of a total of 19 (panel A) and 5 (panel B) questions which related to each of the items on their relevance to assess developmental delayed children in Hong Kong. "Relevance" referred to the appropriateness of its elements for the targeted construct and function of assessment (Haynes *et al.*, 1995). A typical example of the question was "How well is this item relevant to the assessment of the language ability of the developmental delayed children?" The panel members were requested to rate this statement by a four point Likert scale ranged from 1 (poor); 2 (fair); 3 (good); to 4 (excellent). They were also asked to write their comments following each question especially to those which they rated "fair" or "poor". In

addition, the panel members were asked to evaluate the representativeness of all the items under the same sub-scale in assessing the related function of the child. "Representativeness" referred to the degree to which its elements are proportional to the facets of the targeted construct (Haynes *et al.*, 1995). A typical example of a question was "How well do the four test items listed above completely represent the assessment of the language function of the developmental delayed children in Hong Kong?" The same four point Likert scale was used and written comments were requested from the panel members.

Method of Data Analysis

The modes of the ratings of each item in the questionnaires were calculated. The percentages of agreement of the mode ratings among all panel members were computed. As the questions utilized a four point Likert scale, a mode score at or above 3 indicated that most of the panel members' comments on the "relevance and representativeness" of the items were "good" or "excellent" whereas a mode score at or below 2 indicated that most of the panel members' comments were "fair" or "poor". The consistency of the mode score could be reflected by the percentage of agreement. The closer the agreement was to 100%, the more consistent comments were from the panel members. In addition, all questions which were rated at or below 2 by any of the panel members were discussed among the members to reach a compromised decision on retention or modification of the translation.

Phase IV: Preliminary Field test to Language Items

The aim of this phase was to explore the difficulty level, the fluency and clarity of the translated and the newly designed questions in the four language specific items: articulation, general information, follow direction and sentence repetition. This was an additional field test specially arranged by the researcher to assess the difficulty level of the questions. Difficulty level of a question or item was defined in terms of the percentage (or proportion) of people who answered correctly. In the process of test construction, the difficulty level enabled the researcher to choose questions of the most suitable difficulty level for the items (Anastasi & Urbina, 1997). The additional field test thus aimed to test the newly designed questions under the language specific items and choose the most suitable difficulty level ones for the final field test.

Principles in Translation of Test Contents

General information

There are three questions in general information item of the original MAP. Altogether six questions were prepared in the preliminary field test. The translation of this item mainly followed those of the original version. In addition to the three questions by direct translation from the respective age groups, the three questions of the other age groups were also added. For instance, age group II included questions in age group II and IV. For those in age group VI, besides answering questions of their age group, they also answered three questions designed by the researcher. These three questions were chosen from the vocabulary list in the Reynell Developmental Language Scale (Chinese version) (Ko, *et al.*, 1987) to same age group of subjects.

Follow Direction

There are also three questions in the follow direction item of original MAP. A total of six questions were prepared for preliminary field test. The questions requested the subjects to follow instructions of the tester with the tested materials provided. The style of assessment was very similar to that of the Reynell Developmental Language Scale - Comprehension A in which the child was also requested to follow instructions. The Reynell Developmental Language Scale (RDLS) was developed by Reynell (1975) and revised in 1985 by Reynell and Huntley. There were test items designed to assess the children's verbal comprehension and verbal expression. The age range of the test was up to seven years old. The RDLS was translated the test into Cantonese and norm from about 2000 children in Hong Kong was taken (Ko, *et al.*, 1987). It specified some of the differences in language structure between Cantonese and English including the use of vocabulary, localizers and classifiers. In the tense structure, there was no formal tense structure in Cantonese. Cantonese used the perfectives in expressing the timing of the events.

New questions based on the RDLS were designed with consideration to the available materials in original MAP. Additional materials were minimal. Only one cup (for assessment of preposition) and two pencils with different length (for assessment of adjective) were added in the preliminary field test.

Sentence Repetition

Sentence Repetition Assessment was based on the concept of Hewlett's model of phonological processing and phonetic production (Hewlett, 1990). Hewlett (1990) proposed that there were two lexicons: input and output. The input lexicon contained perceptual representations in terms of auditory-

perceptual features while the output lexicon stored the articulatory representation of words. The realization rules mapped the perceptual representations onto articulatory representations. The motor programmer received the auditory-perceptual representation of a word and attempted to devise a motor plan for its production. The motor processing component assembled the motor plan of the sequence of gestures involved in pronunciation of the word, and determined the precise values of the articulatory parameters involved. In the case of sentence repetition, the word was selected directly from the output lexicon and implemented by highly learned combinations of muscle commands (Ng, 1995). It required a higher-order semantactic category , owing to the delay in saying the target words, the children were not purely imitating, but may be considered as part of their spontaneous speech mode (Maxwell & Rockman, 1984). Thus, the sentences designed for sentence repetition test have to be familiar and age appropriate to Hong Kong children in their respective age range.

Similar to the other two items, there were three questions in the original MAP. A total of nine questions, six newly designed and three by direct translation prepared for the preliminary field test. As required by the original MAP, the child must repeat the sentences in correct order and grammar. Credit was not given if the child omitted one or more words, substitues words, or inserts extra words.

The complexity of the sentences in different age groups increased by increasing the number of words in the sentence, the difficulty level of the vocabulary, localizers, classifiers and perfectives used.

Articulation

The articulation item was assessed by asking the subjects to repeat about 36 phonemes in English in the original MAP. In Cantonese, the pronunciations of the phonemes involves two main components besides the vowels: initials (onsets) and finals (codas). So (1993) had conducted an extensive research on the children's developmental age in acquiring these phonemes. Thus, the articulation item in CMAP mainly follows this guideline in designing a total of 19 onsets and 8 codas, each assessed twice by two words for further selection.

Preliminary Field Test : Stage One

Sample

Altogether sixty normal children were selected by convenience sampling from two kindergartens. The telephone directory was used to find the names of the two kindergartens. The person in-charge of each kindergarten was then contacted. The researcher explained the purpose of the field test and the data collection procedures. The first two kindergartens contacted according to the telephone directory, whose person in-charge showed interested in the research, were invited to join this field test. Ten children from age group II, IV and VI (according to Table 2.1) were selected randomly from each kindergarten. All children's medical report to the kindergarten and also the result of their developmental screening assessment at the Department of Health before age three were reviewed to rule out children who were previously diagnosed as developmental delay or having other medical problems. The inclusion and exclusion criteria of subjects were the same as that in phase VI.

Method of Data Collection

The researcher sent consent letters, through the kindergartens, to parents of these children one week before the administration of the test. All the selected children were invited to perform a 20 minute test on the four language specific items (Appendix F) in a room provided and conducted by the researcher. The children who refused to do the test may leave at any time. Souvenirs were given to those who joined the test.

Data Analysis

For items: general information, follow direction and sentence repetition which consisted of three questions each, questions designed are screened with respect to their difficulty level at stage one. According to original MAP, the questions in which at least 75% of normal children should be able to answer them correctly. Thus, the questions with difficulty level below 0.75 were eliminated for preliminary field test at stage two. For articulation item, all the phonemes in onsets and codas were assessed especially on the practicability of the use of codas in the assessment.

Preliminary Field Test : Stage Two

Sample

A total of sixty subjects were selected by cluster random sampling to ten kindergartens, with twenty subjects from each age group. Ten districts were randomly selected from the eighteen districts of Hong Kong. One kindergarten were selected from each district. The selection criteria was the same as that in stage one.

Method of Data Collection

The procedures were the same as that in stage one.

Data analysis

Difficulty levels of all the questions in the items were computed. According to the percentile scoring system of the MAP, 0 – 5th percentile represented children with developmental delay; 6th – 25th percentile represented children at risk and above 26th percentile implied that the children were normal. In original MAP, the number of correct answers to the questions was the raw score. It was then converted to the percentile scores mentioned above. Different language items in different age group had their unique difficulty level. Thus, the difficulty level of the selected questions for the language item also followed the principle of MAP. The combination of three were all tested so as to choose the best combination which fitted the percentile score and the raw scores of the MAP in its respective age groups.

Phase V: Pilot Test of CMAP

The objective of this phase was to evaluate the practicability of the CMAP for children in Hong Kong. The practicability includes the understandability of the instructions by the children and the acceptability of the time limits in practice. It allowed the researcher to learn the potential problems faced by those responding to the CMAP (Geisinger, 1994). These potential problems included the actual set up time in a kindergarten setting, the technical support required from the kindergarten, the children's general response to the CMAP, the administration time required, and the clarity of the instruction sheet to the research assistants.

Sampling

One of the kindergartens which joined the preliminary language field test in phase IV was contacted again to invite the children to join this pilot test. Altogether, nine children, three from each age group (II, IV and VI) were chosen randomly to do the finalized CMAP. The procedure for collecting consent from their parents was similar to phase IV. The children completed the 45 minute CMAP test in the room provided by the kindergarten. Souvenirs were given to them after they completed the test. Each of the three research assistants conducted the CMAP on three children, one from each age group. Throughout the whole process video taping was performed. The researcher was in-charge of the entire event and supervised the research assistants during the data collection.

Method of Data Collection and Data Analysis

The researcher held a three hour meeting with the three research assistants. The data collected from the field test was analyzed. The research assistants were requested to comment on the actual set up time required in the kindergarten, the technical support required from the kindergartens, the children's general response to the CMAP, the administration time required, and the clarity of the instruction sheet to the research assistants. Video play back was prepared to clarify some of the practical problems of the research assistants during the pilot test. A list of procedures in setting up and the technical requirement of the kindergartens was made for the final field test.

Phase VI: Construct and Structural validity of CMAP

Introduction

The aim of phase VI was to evaluate the construct and structural validity of the three sub-scales of the CMAP by factor analysis, assessment of the internal consistency of the items in each sub-scale and the item distribution of the items by the experimental testing of the field test (Anastasi & Urbina, 1997). In addition, the gender difference in performance of the three sub-scales of the CMAP were examined.

During the field test, the performance of the children in Hong Kong in the items under the foundations, coordination and verbal sub-scales were assessed. These children were selected by cluster random sampling. The gender difference was evaluated by comparing the children's performance between male and female subjects within the same age group. Moreover, the structure was reviewed by evaluating the items in relation to the three sub-scales by factor analysis, inter-item correlation, internal consistency and item difficulty and discrimination.

Sample

The subjects selected were normal children attending kindergartens or nurseries in Hong Kong. The kindergartens were selected by cluster random sampling method. This was because it was not practical to obtain a complete listing of all preschool children among the whole population. By adopting the cluster sampling approach to link the preschool children to the "already established group" (Portney & Watkins, 1993) that is, the kindergartens and nurseries within the districts, the subjects could be contacted. Besides, the sampling error was minimized. The sampling was multi-staged. The first stage

was the selection of sample districts in Hong Kong (area probability sampling). From the eighteen districts of Hong Kong (defined by the District Board of Hong Kong), ten districts were randomly selected. The method of selection was by putting all the names of the districts on a equal size piece of paper into a box. Ten pieces of paper were then drawn out. The second stage was the selection of kindergartens/nurseries in the selected districts. All the names of the kindergartens and nurseries in the ten selected district were put into a box for random sampling. One kindergarten or nursery was randomly selected from each district. All the chosen samples were contacted by phone and letter to explain the aims of the research. For those who refused to take part in the research, the names were put back into the box and another kindergarten was chosen for replacement. A total of ten kindergartens were selected.

After confirmation with the person in-charge of the kindergartens, the kindergartens were requested to prepare a list of children who fell within the selected age ranges. The field test selected subjects from three age groups among the six age groups in MAP. Children with age ranges three years three months to three years eight months(Age Group II); four years three months to four years eight months (Age Group IV); and five years three months to five years eight months(Age Group VI) were selected. Altogether, three female and three male subjects in each of the three age groups were randomly selected from each kindergarten. This was done by requesting the kindergarten's in-charge to select children randomly from each age group.

Inclusion and Exclusion Criteria

The children's parents were contacted for the consent and ask for background information on the development of their children. Children who

had physical problems (such as cerebral palsy) or delayed development, found during their assessment at the Comprehensive Observation Scheme in the Health Centres of Department of Health, were excluded from this research. The Comprehensive Observation Scheme was a free of charge screening assessment provided by the Hong Kong SAR Government for children from age zero to age three. Children in Hong Kong are invited to join this Scheme to screen for developmental delays. The parents of the selected subjects had to fill in a questionnaire to review their children's performance during the Comprehensive Observation Scheme. In addition, the medical reports of the children when they entered the kindergarten or nurseries were also checked. This was done with parents' consent and was to confirm that the children were apparently normal during the data collection period.

There was a total of eighteen subjects selected from each kindergarten. As there were ten kindergartens taking part in the field test, the total number of subjects involved was 184.

Method of Data Collection

Consent from parents of selected subjects

Through the kindergartens, letters and written consents (Appendix G), for participating in the research, was sent to the parents of selected children. The letter explained the nature and purpose of the research. To ensure that the child selected fulfilled the inclusion and exclusion criteria, parents were asked to fill in a data sheet on the child's health history and their consent to allow us to screen the medical check up report before entering school. For those who refused to participate, a replacement subject was randomly selected. All the

subjects had written consent from the parents before actual administration of the test.

Testers for administration of the Test

The research aimed at studying about 180 subjects by visiting ten to eleven kindergartens. Three research assistants were involved to speed up the data collection time. These research assistants were recruited from final year Occupational Therapy students at the Hong Kong Polytechnic University. They all had experience in working with children during their previous clinical placement. They attended a training workshop run by the researcher. It included explanation of the aims of the CMAP, demonstration of the procedures in carrying out the items, filling in the data sheet and questioning time for any clarification needed. They tried out the CMAP during two three hour practical sessions under the supervision of the researcher and applied the CMAP to a total of six children.

Inter-rater reliability of the testers

The purpose of the inter-rater reliability test for the research assistants was to ensure the consistency of the obtained scores among the three research assistants and the researcher.

Sampling - Sixteen subjects were selected by convenience sampling from one of the kindergartens which had agreed to join the field test at phase VI. The subjects were evenly selected from different gender and age groups.

Data Collection - One research assistant conducted the CMAP assessment on the subjects and was video taped by the researcher. The researcher and the three research assistants independently rated the children's performances by watching the video playback.

Data Analysis - Since the scores in most of the items were ordinal data such as: able to do two out of three tasks assigned, etc., the correlation coefficient was calculated by Spearman correlation to all the ratings of the three research assistants and the researcher.

Schedule of Administration of the Test

By referring to the experience in the pilot test, the actual administration time of the CMAP was about 45 minutes for each subject. To assess 18 subjects in each kindergarten or nursery with three research assistants together, it took about four and a half hours. Instead of taking the data, the researcher coordinated and supervised the ongoing data collection process. This was to ensure that the supervision was close and intensive. The research assistants were staying in the same area while carrying out the data collection. By following the usual time table of the kindergartens, the children were available after morning greeting and before lunch. This would be about two and a half hours. Thus, two half day sessions were planned in each kindergarten. The actual date and time was scheduled with the kindergartens. Since the children selected were mainly attending half-day school, most of the data collection was done within one morning/afternoon session or over two days.

The data collection started in late June till mid July. This was the best possible time for the research assistants as they could have one month training after their final examination. On the other hand, the children had had their final examination and were waiting for the summer holidays. This was also an ideal time for the children of age group II (three years three months to eight months) as they were more familiar with the kindergarten and had almost

finished the school term. They were easier to approach during the test while they were accompanied by their teachers. In order to minimize stranger anxiety, the research assistants and the researcher played with the children for a while to warm up prior to administration of the test. The teacher of the child may stay with the child provided they do not give hints to the child during the test. If the child refused to do the test, he/she could leave at any time.

Souvenirs

Souvenirs were given to the children whether they successfully completed the test or not. All test content followed exactly the same as the MAP except the translated instructions and the language specific items. A thank you letter and card was sent to the kindergartens that participated in the research. Consent from the teachers was obtained before distributing the souvenir to the children. Since some of the parents requested the results, simple result sheets were given to them via the in-charge of the kindergartens or nurseries. The result sheets had stated that all the results were calculated by the MAP United States version only. As some of the parents were very keen in joining the study, the actual number of subjects joining the field test was slightly more on the request of the person in-charge from the kindergartens or nurseries.

Time table in data collection

9:00 - 9:05 am - Met the school in-charge

- Obtained the name list of all the subjects with their date of birth, consent and medical information

9:05 - 9:15 am - Set up the video equipment, furniture and gross motor items and screen the informations

9:15 - 11:55 am - Led the first group of children to the testing area

- Matched name, date of birth and age group of the subject
- Conducted the test on the subjects
- Presented the souvenir to the subjects

11:55 - 12:00 pm - Presented the thank you letter to the in-charge

Method of Data analysis

Item Distribution Examination

The descriptive statistics of the raw scores of the subjects' performance were assessed. The purpose was to examine the distribution of the raw scores of the individual items and identify those with zero variance or exceptional skewness and kurtosis.

The procedure began firstly with input of the raw scores of the data. Descriptive statistics of all items in individual age groups were examined. Raw scores with zero variance or high kurtosis and skewness were sorted out. Then the polarity of the data were changed where necessary so that the ascending numbers would imply good performance. The range of the raw scores were converted to standard scores of 0.00 to 1.00. This was to ease the comparison with performance of children in United States at later stage.

Comparing the performance between male and female subjects

The purpose was to examine whether there was sex difference in performance in the items. The raw scores of each items among the male and female subjects in individual age groups were compared by Kruskal-Wallis one-way analysis of variance. It was the most appropriate way to handle ordinal level data when more than two groups were compared (Portney & Watkin, 1993). As the raw data of the items were ordinal data, and the performance of the male and female subjects across three age groups were compared, the Kruskal-Wallis one-way analysis of variance was the most suitable method to do the analysis. After initial analysis, as there were 18 comparisons done, adjustment of type I error was made by Bonferroni's Correction (Portney & Watkins, 1993).

Internal consistency and inter-item correlation

The purpose of this step was to identify the sets of variables that linearly correlated with each other under the same sub-scale by creating a correlation matrix of the items within same sub-scale and same age group.

The inter-item correlation within same sub-scale: foundations, coordination and verbal in individual age groups were calculated by the Spearman correlation and the significant levels were also evaluated. Also, the homogeneity of items to the sub-scales were assessed by calculating the coefficient alpha, that is, the internal consistency of the items to the sub-scales (Clark & Watson, 1995).

Item Analysis

The difficulty level of all the 18 items were assessed. Difficulty level of an item was defined in terms of the proportion (or percentage) of people who

answered correctly (Anastasi & Urbina, 1997). The item discriminative index were also calculated by adding the total item standard scores and correlating it to each of the 18 items.

Factor Analysis

The purpose of factor analysis was to analyze the interrelationships of the items to the related sub-scales by evaluating the factorial composition of the items, together with the loading of each major factor and the correlation of the test with each factor (Anastasi & Urbina, 1997; Clark & Watson, 1995; Smith and McCarthy, 1995). Exploratory factor analysis was adopted to condense the items to their related sub-scales (Nunnally & Bernstein, 1994).

Exploratory factor analysis involved: 1) subject the items to a principal components analysis (Comrey, 1988), and extract the first factors; 2) examine the loading of the items to the factors. Items that loaded relatively strong on one factor and weak on other factors were excellent items within the sub-scale (Clark & Watson, 1995).

Factoring began by computing a correlation matrix and observing the groupings in the data, the signs and sizes within groupings, and investigating the correlation between groupings. Then, one had to create a rotated factor matrix to confirm statistically which items related highly to which unique sub-scale (Nunnally & Bernstein, 1994). Scree plot of successive eigenvalues against their ordinal positions were examined to reduce the number of factors where necessary. Finally, the researcher named the factors and matched them with the original version of MAP.

CHAPTER IV

RESULT

Introduction

Chapter IV describes the translation process and results in testing of equivalence of the translated instructions for the foundations, coordination, and verbal sub-scales; test of equivalence of the translated items; evaluation of content validity; the inter-rater reliability estimates of the translated version; difference in performance on CMAP between male and female children and item distribution, inter-item correlation and internal consistency and factor structure of the three translated sub-scales.

Translation and Test of Equivalence to Instructions of Items

Translation and Backward Translation

The instructions of all 18 items of MAP were translated to Cantonese by a qualified translator and a backward translation of Cantonese instructions to English was done by another independent qualified translator. The procedures followed the protocol in translating standardized instruments specified by Satorius and Kuyken (1994). The experience gained in the backward translation suggested three main areas of discrepancy between the original version and the backward translated version.

Firstly, there were occasions when the translated English version did not use the exact wordings but words which had similar semantic meanings. For example: "We are going to switch hands" (original) when compared with "Let us use the other hand for a change (translated); show me with this hand"

(original) when compared with "point it out with this hand" (translated). Secondly, there were two sentences which had inappropriate elaboration of the original sentence in the translated version. For example: "make it as big as you can" (original) which appeared as "make it as high as possible; the higher, the better (translated). Thirdly, in some cases, the differences between the original and translated versions were attributed to differences in grammatical structure of the English and Chinese language. These differences disturbed the fluency of the translation. It will be tackled during the test of equivalence by the expert panel reviews in the next stage.

Test of Equivalence

Two parallel sessions of panel review were held to evaluate on the degree of equivalence of the CMAP after backward translation. Panel A consisted of six occupational therapists. The average years of experience of the members in working with children was 8.3 years. Panel B consisted of six speech therapists and their average years of experience was 5.2 years. Both panels were requested to review the fluency and semantic equivalence between the original and translated versions. Panel A reviewed 14 items mainly in the foundations and coordination sub-scales while panel B reviewed the 4 language specific items.

In terms of "fluency", the panel members agreed that items were appropriately translated in 13 out of 18 items with high average percentage of agreement of 79% (range: 67 - 100%). Four items: stepping, walks line, rapid alternate movement, and digit repetition were rated as "disagree" (2) rating with relatively low percentages of agreement (50%). In these items, 50% of the panel members either rated "agree" (3) or "strongly agree" (4) while the other

50% rated "disagree" (2). The motor accuracy item received a higher consistent rating of "disagree" (2) (83% of agreement) from the panel members.

In terms of "semantic equivalence", the panel members either rated "agree" (3) or "strongly agree" (4) in 17 out of 18 items which meant similar semantic meaning was retained after the translation. The average percentage of agreement was 81% (range: 67 - 100%). Only the stepping item received either a disagreed (2) rating (50% of agreement) or an agreed (3) rating (50% of agreement).

Detailed descriptions of the panel members' ratings on fluency and semantic meaning are shown in Table 4.1. Comments made by panel members who disagreed (either 1 or 2 ratings) were collected for further discussions by the panel. By discussion, a consensus decision was obtained among the panel members. Modifications to the testing instructions were then made to motor accuracy, stepping, walks line, rapid alternate movement and digit repetition items to improve their fluency. In addition, modification was made to the stepping item on its semantic meaning. Detailed descriptions on the modification made to individual items can be found in Appendix H.

Table 4.1

Degree of Equivalence between the MAP and CMAP as rated by Panels A & B

Items	mode rating score (% of agreement)	
	fluency	semantic meaning
Tower	3 (100%)	3 (67%)
Stereognosis	3 (83%)	3 (83%)
Finger localization	3 (83%)	3 (67%)
Motor accuracy*	2 (83%)	3 (100%)
Vertical writing	3 (67%)	3 (67%)
Hand-to-nose	3 (67%)	3 (67%)
Romberg	3 (83%)	3 (100%)
Stepping*	2/3 (50%)	2/3 (50%)
Walks line*	2/3 (50%)	3 (83%)
Supine flexion	3 (67%)	3 (100%)
Kneel-stand	3/4 (50%)	3/4 (50%)
Tongue movement	3 (83%)	3 (83%)
Rapid alternate movement*	2 (50%)	3 (83%)
General information	3 (67%)	3 (83%)
Articulation	3 (100%)	3 (100%)
Follow direction	3 (83%)	3 (83%)
Sentence repetition	3 (83%)	3 (83%)
Digit repetition*	2 (50%)	3 (83%)

Rating Scale: 1 - strongly disagree; 2 - disagree; 3 - agree; 4 - strongly agree

*Modification made to items

Although the semantic meaning of the CMAP after backward translation being regarded as satisfactory, the fluency and incompatible grammatical structure between English and Cantonese still presented with major problems. The panels found that the fluency and understandability of some instructions were not satisfactory. As a consequence, 13 out of 18 items' instructions underwent modifications. For example: "by accident" originally translated to "意外" (yee oi) was changed to "唔小心" (ng siu sum) in Cantonese and "move your tongue up" originally translated to "郁" (yuk) was changed to "伸" (sun). These changes made the instructions more practical and applicable to the children in Hong Kong.

Translation and Evaluation of Equivalence to Four Language Specific Items

Test of Equivalence

Panel B in phase I also evaluated the equivalence of items: articulation, follow direction, general information and sentence repetition between the CMAP and original MAP. The mode rating score was "agreed" (3) that the fluency and semantic equivalence were maintained in the translated version of articulation, follow direction and sentence repetition items (range: 83 -100% of agreement). There were two panel members who suggested rephrasing two of the questions in the general information item for age group IV. The suggestions were to use two appropriate classifiers to describe "an orange" as "個" (goh) instead of "隻" (jek) and describe "how many pieces" as "舊" (kau) instead of "塊" (faai). After an overall discussion among all panel members, the

suggestions were accepted and the translated version was thus modified accordingly.

Content Validity of CMAP

Introduction

The same expert panels A and B were involved in evaluating content validity of the three sub-scales of CMAP. Panel A evaluated items under the foundations and coordination sub-scales while panel B evaluated items under the verbal sub-scale.

Relevance of CMAP

Among the items categorized under foundations sub-scale (10 items), coordination sub-scale (7 items) and verbal sub-scale (4 items), the mode scores were assigned by the panel members either "good" (3) or "excellent" (4) (Table 4.2). These reflected that a majority of panel members agreed all the items were relevant to assess children's developmental delay in their respective areas. However, the percentages of agreement of those ratings ranged from low to high (50% to 83%). Inconsistent ratings were found in the tongue movement and rapid alternate movement items under the coordination sub-scale (50% of agreement). The lower consistency was found to be due to half of the panel members rating "excellent" (4) while the other half rated "good" (3) on the tongue movement item. In contrast, for the rapid alternate movement item, panel members rated "good" (3) or "excellent" (4) (67% of agreement) and 33% regarded the item as irrelevant, rating "fair" (2) or "poor" (1). Some panel members commented that the "hand-to-nose" (13%) and "kneel-stand" (33%) items' difficulty level was too low.

For the relevance of items under verbal sub-scale, panel members commented that the "n" and "ŋ" onsets under the articulation item were often wrongly pronounced among the Hong Kong people. They suggested deleting it from the articulation item. In addition, the experts in panel B suggested having a preliminary field test of the newly designed questions and the questions directly translated from MAP for the follow direction, general information and sentence repetition items so as to pre-test the difficulty level of these questions before the final field test.

Table 4.2

Results Obtained from the Panels' Comment on "Relevance" of the CMAP

CMAP sub-scales	Items	mode score (% of agreement)
Foundations	stereognosis	3 (83%)
	finger localization	3 (67%)
	vertical writing	3 (67%)
	hand-to-nose	3 (83%)
	Romberg	3 (67%)
	stepping	3 (83%)
	walks line	3 (83%)
	supine flexion	3 (67%)
	kneel-stand	3 (67%)
	rapid alternate movement	4 (50%)
Coordination	tower	4 (67%)
	motor accuracy	3 (67%)
	vertical writing	3 (67%)
	walks line	3 (83%)
	tongue movement	3/4 (50%)
	rapid alternate movement	4 (50%)
	articulation	3 (67%)

(continue over)

Table 4.2 (Continue)

Results obtained from the Panels' comment on "relevance" of the CMAP

CMAP sub-scales	Items	mode score (% of agreement)
Verbal	general information	3 (83%)
	follow direction	3 (83%)
	sentence repetition	3 (83%)
	digit repetition	3 (67%)

Rating Scale : 1 - poor; 2 - fair; 3 - good; 4 - excellent

*Shaded areas are those with lower than 67% of agreement

Representativeness of CMAP

As for representativeness of the items, all panel members rated either "good" (3) (67%) or "excellent" (4) (33%) in the ten items covering the foundations sub-scale. Similarly, all of the panel members rated good (3) (87%) or excellent (4) (13%) in the seven items covered the coordination sub-scale. For the verbal sub-scale, all panel members rated good (3) (100%) in the four items of this sub-scale.

Preliminary Field Test on Three Language Items

The purpose of the preliminary field test was to explore the difficulty level of the translated and the newly designed questions (63 questions) in general information, follow direction and sentence repetition items. In the general information and follow direction items, there were three questions from the direct translation of the MAP and three newly designed questions. In the sentence repetition item, there were three questions from the direct translation of the MAP and six newly designed questions. The newly designed questions

had been reviewed by the panel. The panel members commented that limited research had been done in these areas, thus, the questions should be assessed in the preliminary field test. These questions were screened by 60 children (of the same age range as this field test) from two kindergartens chosen by convenience sampling in Shatin and the Kowloon East Area in stage one of preliminary field test. The questions with difficulty levels below 0.75 (as MAP suggested) were deleted. A total of four to eight questions were left for the preliminary field test stage two. Meanwhile, in the articulation item, 19 onsets and 8 codas were also tested in stage one of preliminary field test. As the onsets: "n" and "ŋ" were very difficult for the children (difficulty level: 0.25), they were permanently deleted from the preliminary field test stage two. In addition, the codas were also eliminated due to the low consistency found among the testers obtained during preliminary field test stage one. Therefore, in the articulation item, 17 onsets were retained for the preliminary field test stage two.

Demographic Characteristics of the Children

A total of sixty Chinese children were selected randomly from ten kindergartens from ten districts of Hong Kong. Twenty children from age group II, IV, and VI respectively were randomly selected. The mean age and gender distributions are described in Table 4.3.

Table 4.3

Demographic Characteristics of Subjects in Preliminary Field Test (stage two)

Age Group	Male	Female	Mean Age
II	10	10	3.4 years
IV	10	10	4.5 years
VI	10	10	5.4 years

Results in Item Analysis

The difficulty level of the directly translated and newly designed questions for the three language specific items: general information (4 - 5 questions), follow direction (4 - 5 questions), and sentence repetition (7 - 8 questions), was evaluated in terms of the item difficulty (proportion of subjects who answered correctly) (Anastasi & Urbina, 1997). The difficulty levels of all the questions were computed. The choice of appropriate item difficulties varied with the purpose of the test construct (Anastasi & Urbina, 1997). In MAP, three questions were designed for each language item. Different age groups might have different questions. The children's performance in each item was determined by the number of questions answered correctly which is converted to percentile scores (below 5^{th}; between 6^{th} and 25^{th}; and above 25^{th} percentile). For example, for general information item, answering zero question had 5^{th} percentile score, and one question had 25^{th} percentile score. Thus, item difficulty of all the combinations of any three questions in each items were tested to select the best three questions which fitted the percentile

score of the questions in MAP in each age group. The difficulty level is listed in Table 4.4, 4.5 and 4.6.

Table 4.4

Difficulty Level of General Information Item

Questions	Item Difficulty Level (0.00 - 1.00)		
	Age Group II (n = 20)	Age Group IV (n = 20)	Age Group VI (n = 20)
Q1	0.90	0.97	0.98
Q2	0.92	0.73	0.65
Q3	*0.30*	0.93	0.89
Q4	*0.95*	*0.95*	*0.58*
Q5	*0.88*	*0.44*	*0.52*

Table 4.5

Difficulty Level of Follow Direction Item

Questions	Item Difficulty Level (0.00 - 1.00)		
	Age Group II (n = 20)	Age Group IV (n = 20)	Age Group VI (n = 20)
Q1	*0.70*	*0.70*	0.86
Q2	0.93	*0.57*	*0.75*
Q3	*0.86*	0.85	*0.62*
Q4	0.67	0.67	*0.86*
Q5	*0.82*		*0.92*

Remarks: In Table 4.4, 4.5 and 4.6, figures in italics are newly designed questions while others are original questions

Table 4.6

Difficulty Level of Sentence Repetition Item

Questions	Item Difficulty Level (0.00 - 1.00)		
	Age Group II (n = 20)	Age Group IV (n = 20)	Age Group VI (n = 20)
Q1	*1.00*	*0.70*	*0.66*
Q2	*0.90*	*0.97*	*0.77*
Q3	*0.95*	*0.60*	*0.40*
Q4	*0.93*	*0.90*	*0.92*
Q5	0.97	*0.85*	*0.62*
Q6	0.82	0.82	*0.91*
Q7	0.87	0.62	0.54
Q8			0.65

After the total scores of all combinations of three questions under the respective age groups were made, the scores were matched with the original percentile scores of MAP. Three questions were selected for each of the three items. Example is shown in Table 4.7.

Table 4.7

Question Selection in Follow Direction Item at Age Group II

Combinations	0 - 5th percentile	6 - 25th percentile	above 25thpercentile
original MAP	-	0 - 2	3
Q1, Q2, Q3	-	0 - 2	3
Q1, Q2, Q4	0 - 1	2	3

In this case, Q1, Q2 and Q3 combination was selected. In age group IV, since a perfect match was not found, the best possible match was selected (Table 4.8).

Table 4.8

List of Selected Questions for Items:

general information, follow direction and sentence repetition

Items	Questions		
	Age Group II	Age Group IV	Age Group VI
general information	Q1, Q3, Q4	Q1, Q2, Q3*	Q3, Q4, Q5
follow direction	Q1, Q2, Q3	Q2, Q3, Q4*	Q1, Q2, Q3
sentence repetition	Q1, Q2, Q4	Q3, Q4, Q5	Q1, Q2, Q5

* Questions were not the perfect match

Inter-rater Reliability of CMAP

Before the final field test, the inter-rater reliability among the researcher and the three research assistants was estimated. Altogether 16 subjects were recruited from the three age groups II, IV and VI. The distribution was: three male and three female subjects in age group II; two male and two female subjects in age group IV and three male and three female subjects in age group VI.

The inter-rater reliability was computed by the Spearman rank correlation coefficient since the scores of the items were ordinal data (Portney & Watkins, 1993). Among the 18 items, all except the rapid alternate movement and Romberg items had very satisfactory ratings. The reliability coefficient $r \geq 0.90$ ($p < 0.01$). The items: hand-to-nose, kneel-stand, general information, follow

direction and sentence repetition yielded zero variance in the inter-rater scoring of children's performance. It implied that 100% agreement between all the raters was found. The Romberg item had lower but good reliability coefficient r = 0.80 to 0.99 ($p < 0.05$) while the rapid alternate movement item had moderate reliability coefficient r = 0.62 to 0.70. Detailed description on the inter-rater reliability is listed in Table 4.9.

The rating method of the rapid alternate movement item was re-examined. There were some discrepancies among the research assistants in counting the steps made by the subjects. An extra practice session was arranged for the three research assistants. The inter-rater reliability on this item was tested again on seven subjects from the first kindergarten that joined the research. The results looked much more promising. The correlation coefficient r = 0.83 to 0.94 ($p < 0.001$).

Table 4.9

Inter-rater Reliability of CMAP in Three Sub-scales

Items	Reliability Coefficient
Tower	1.00***
Stereognosis	1.00***
Finger localization	1.00***
Motor accuracy	1.00***
Vertical writing	0.93 - 0.99***
Hand-to-nose	zero variance
Romberg	0.80 - 0.99**
Stepping (distance/deviation)	1.00***/1.00***
Walks line (speed/error)	0.99 - 0.99***/0.82 - 1.00**
Supine flexion	0.90 - 0.99***
Kneel-stand	zero variance
Tongue movement	0.98 - 0.99***
Rapid alternate movement	**0.62 - 0.70****
General information	zero variance
Articulation	0.93 - 0.97***
Follow direction	zero variance
Sentence repetition	zero variance
Digit repetition	1.00***

*$p < 0.05$. **$p < 0.01$ ***$p < 0.001$

Construct Validity of CMAP in Three Sub-scales: Foundations, Coordination and Verbal

Sampling

The subjects were selected from ten different districts. One kindergarten was selected from each of the ten districts. Altogether ten kindergartens had participated in the research. In each kindergarten, 14 to 22 children participated in the research with an average of 18 children. Their medical report and parent report showed that no abnormalities were found during the developmental screening assessment done by the Department of Health before age three. The subjects were all Chinese. The demographic characteristics of the subjects are listed in Table 4.10.

Table 4.10

Demographic Characteristics of Subjects

Age Group	Male	Female	Mean Age
II	30	31	3.4 years
IV	29	31	4.5 years
VI	32	32	5.5 years

Although the MAP did not suggest gender difference in performance, the difference in performance between male and female subjects of CMAP was reconfirmed by using the Kruskal-Wallis one way analysis of variance. Among the 18 items tested, there were three items which found significant differences between male and female children after the initial analysis. They were the

items: tongue movement in subjects of age group II and IV. The mean difference was 0.48 (chi square = 3.90, p = 0.05) and 0.98 (chi square = 4.70, p = 0.03) respectively. In the item stereognosis at age group II, the mean difference was 0.37 (chi square = 5.06, p=0.02) and in item tower at age group VI, the mean difference was 0.13 (chi square = 1.73, p=0.04). However, as the initial analysis involved 18 items, after adjusting the type I error with Bonferroni's correction (0.05/18), the new actual p value became 0.0028 (Portney & Watkin, 1993). Thus, no significant difference in item scores were found between male and female children in all items. Hence, all subsequent analysis was carried out by pooling both male and female subjects.

Performance of the Children in Hong Kong

Item Distribution

The results showed that there were two items in which subjects gained the highest score with zero variance. These items included Romberg at age group IV and kneel-stand at age group IV and VI. In addition, there were items in which most of the subjects gained very high score. These included stereognosis at age group II and IV, kneel stand at age group II, hand-to-nose at age group II, IV and VI, and Romberg at age group II. Their mean scores and standard deviations are included in Table 4.11.

Table 4.11

Performance of Children in Hong Kong in CMAP

Items (range)	Mean Score (SD)		
	Age Group		
	II	IV	VI
Tower (0-16)	11.13 (2.76)	11.68 (2.75)	12.84 (2.29)
Stereognosis (0-4)	**3.93 (0.25)**	**3.93 (0.25)**	3.67 (0.62)
Finger localization (0-4)	2.00 (1.21)	2.68 (0.91)	3.41 (0.68)
Motor accuracy (0-18)	5.77 (3.40)	8.83 (3.40)	10.52 (2.99)
Vertical writing* (0-14)	5.95 (3.06)	6.25 (3.05)	4.31 (2.14)
Hand-to-nose (0-6)	**5.74 (1.05)**	**5.93 (0.52)**	**5.98 (0.12)**
Romberg (0-15)	**14.84 (0.82)**	**15.00 (0.00)**	13.50 (3.20)
Stepping (distance)*(0-5)	1.20 (0.57)	1.20 (0.61)	1.17 (0.58)
Stepping (deviation)*(0-5)	1.95 (1.01)	1.67 (0.95)	1.53 (0.71)
Walks line (speed) (0-39)	11.33 (4.61)	10.92 (6.02)	28.73 (8.93)
Walks line (error) (0-5)	0.97 (1.29)	0.47 (1.69)	2.69 (3.17)
Supine flexion (0-15)	7.59 (4.90)	11.68 (4.64)	13.86 (3.08)
Kneel-stand (0-2)	**1.97 (0.26)**	**2.00 (0.00)**	**2.00 (0.00)**
Tongue movement (0-4)	2.48 (0.91)	3.00 (0.97)	3.38 (0.79)
Rapid alternate movement* (0-15)	8.05 (1.96)	6.83 (1.61)	6.16 (1.28)
Articulation (0-38)	30.46 (2.94)	31.50 (2.90)	32.64 (1.38)

(Continued over)

Table 4.11 (Continue)

Performance of Children in Hong Kong in CMAP

Items (range)	Mean Score (SD)		
	Age Group		
	II	IV	VI
Follow direction (0-3)	2.31 (0.85)	2.13 (0.83)	2.27 (0.84)
General information (0-3)	2.15 (0.68)	2.68 (0.57)	2.03 (0.94)
Sentence repetition (0-3)	2.57 (0.83)	1.90 (0.90)	2.05 (0.86)
Digit repetition (0-5)	2.57 (1.22)	3.67 (0.88)	4.27 (0.70)

*denotes items carrying the higher scores, the poorer the performance

**Items which differ in test content across age groups are shaded

Item analysis of CMAP

The inter-item correlation, total item correlation and the coefficient alpha were computed for all items under the three sub-scales. Items with zero variance or distorted distribution (extreme kurtosis or skewness) were eliminated from the computation as suggested by Clark and Watson (1995).

Inter-item Correlation

According to Clark and Watson (1995), inter-item correlations which fall in the range of .15 to .50 are regarded as possessing good psychometric properties. Results obtained suggested that the inter-item correlations varied across items, sub-scales and age groups.

In the foundations sub-scale, only a few items correlated significantly with each other. At age group II, no items were significantly correlated. At age group IV, items: finger localization and supine flexion (r = .32), and walks line (speed) and rapid alternate movement (r = .29) correlated significantly with

each other ($p < 0.05$). At age group VI, items: stereognosis and stepping (deviation) ($r = .31$), walks line (error) ($r = .41$); and Romberg and walks line (error) ($r = .49$) correlated with each other significantly ($p < 0.05$). For a detailed description, please refer to Table 6.1 to 6.3 in Appendix I. Although more significant results were found in the higher age groups, the number of items involved in the inter-item correlation analysis was lower with lower age groups due to those eliminated items. At age group II, there were five items which reached the ceiling and were deleted for further analysis.

For the coordination sub-scale, similar to the foundations sub-scale, there were no significant correlations found among the seven items for age group II. For age group IV, the items: walks line (speed) correlated significantly with rapid alternate movement and articulation ($r = .29$ and $.25$, $p < 0.05$). Rapid alternate movement and articulation items ($r = .26$) were significantly correlated ($p < 0.05$). For age group VI, even more significant results were found. The tower item correlated with walks line (error) ($r = .32$), tongue movement ($r = .22$) and articulation ($r = .27$)($p < 0.05$). The vertical writing item correlated significantly with rapid alternate movement item ($r = .38$, $p < 0.05$) whilst tongue movement and articulation item also correlated with each other ($r = .26$, $p < 0.05$). For detail description, please refer to Table 6.4 to 6.6 in Appendix I.

In the verbal sub-scale, a comparatively high proportion of the items inter-correlated significantly. For age group II, all the items correlated with each other significantly ($r = .23$ to $.98$, $p < 0.05$). For age group IV, only the general information and follow direction ($r = .49$) and sentence repetition and

digit repetition items (r = .32) significantly correlated ($p < 0.05$). For age group VI, the follow direction item correlated significantly with sentence repetition (r = .33) and digit repetition item (r = .23)($p < 0.05$). Moreover, the general information and sentence repetition items correlated with each other (r = .49, $p < 0.05$). For a detailed description of the inter-correlation, please refer to Table 6.7 to 6.9 in Appendix I.

In conclusion, items in the foundations sub-scale had the lowest inter-item correlation followed by the coordination sub-scale. These two sub-scales presented similar trends of seeing more correlated items in higher age groups. The verbal sub-scale did not have a similar trend.

Total Item Correlation

The total item correlation was calculated in their respective age groups and sub-scales. Items of zero variance were skipped for this analysis. Result reviewed that in foundations sub-scale, most items significantly correlated ($p < 0.05$) with the total score except those with distorted distribution and walks line (speed), walks line (error) and stepping (distance) (Table 4.12). According to Aiken (1997), an item having a correlation close to or less than 0.00 with the total score should be revised or discarded. With reference to data shown in Table 4.12, 50% of items in age group II, IV and 30% of items age group VI need revision.

Table 4.12

Total Item Correlation for Foundations Sub-scale

Items	Age Group II	Age Group IV	Age Group VI
stereognosis	0.15	0.16	0.43**
finger localization	0.38**	0.59**	0.10
vertical writing	0.41**	0.48**	0.40**
hand-to-nose	0.15	0.23	-
Romberg	0.18	-	0.37**
stepping (distance)	0.03	0.18	0.22
stepping (deviation)	0.34**	0.37	0.35**
walks line (speed)	0.14	0.33**	0.04
walks line (error)	0.37**	0.17	0.55**
supine flexion	0.69**	0.54**	0.34*
kneel-stand	0.19	-	-
rapid alternate movement	0.45**	0.36**	0.42**

**$p < 0.01$, *$p < 0.05$

For coordination sub-scale, under same criteria as that in the foundations sub-scale, 62.5% of items at age group II and 25% of items at age group IV and VI should be revised as they are either correlated low with the total item correlation or there were no significant correlation between them (Table 4.13).

Table 4.13

Total Item Correlation for Coordination Sub-scale

Items	Age Group II	Age Group IV	Age Group VI
tower	-0.04	0.24	0.36**
motor accuracy	-0.13	0.50**	0.52**
vertical writing	0.41**	0.53**	0.51**
walks line (speed)	0.14	0.36**	0.44**
walks line (error)	0.37**	0.18	0.12
tongue movement	0.08	0.45**	0.51**
rapid alternate movement	0.45**	0.31**	0.34**
articulation	0.10	0.35**	0.05

**$p < 0.01$, *$p < 0.05$

For verbal sub-scale, the total item correlation ranged from moderate to high (r = .45 - .75, p < 0.01). Relatively lower correlation was found for general information items at age group IV and follow direction item at age group VI (Table 4.14).

Table 4.14

Total Item Correlation for Verbal Sub-scale

Items	Age Group II	Age Group IV	Age Group VI
follow direction	0.68**	0.67**	0.45**
general information	0.61**	0.46**	0.64**
sentence repetition	0.52**	0.66**	0.71**
digit repetition	0.75**	0.58**	0.70**

**p < 0.01, *p < 0.05

Internal Consistency of CMAP Sub-scales for Different Age Groups

The coefficient alpha of the foundations and coordination sub-scales were very low in all three age groups. The verbal sub-scale had moderate to high internal consistency (range r = .49 to .78). Table 4.15 listed the coefficient alpha of three sub-scales in their respective age groups.

Table 4.15

Internal Consistency: Sub-scales versus Age Groups

Sub-scales	Internal Consistency (coefficient alpha)		
	Age Group		
	II	IV	VI
Foundations	0.11	-0.02	-0.09
Coordination	-0.17	0.04	0.30
Verbal	0.78	0.49	0.56

Factorial Structure of CMAP Sub-scales

The factorial structure of items in the sub-scales was explored by factor analysis of different age groups. All 18 items, except those with zero variance and those with distorted distribution (extreme skewness or kurtosis) were included from the analysis. Exploratory factor analysis by principal component methods was employed in the analysis. The eigenvalues of each factor were identified and the scree plot analysis on the eigenvalues was conducted. The number of factors were determined by referring to the scree plot diagram. Factor analysis was then repeated on a specific factor model. Varimax rotation

was finally conducted to strengthen the relation between the variables (in this case, items) and factors (in this case, sub-scales) (Nunnally & Bernstein, 1994).

Factorial Structure of CMAP at Age Group II

For age group II, a total of 11 items (with two measures in stepping item: distance and deviation) were used for the final analysis. A total of 61 subjects were involved in the analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.446 ($p < 0.000$) which indicated that the sample size was adequate for factor analysis. The initial statistics, including the eigenvalues and the percentages of variance of the common factors, are shown in Table 4.16.

Table 4.16

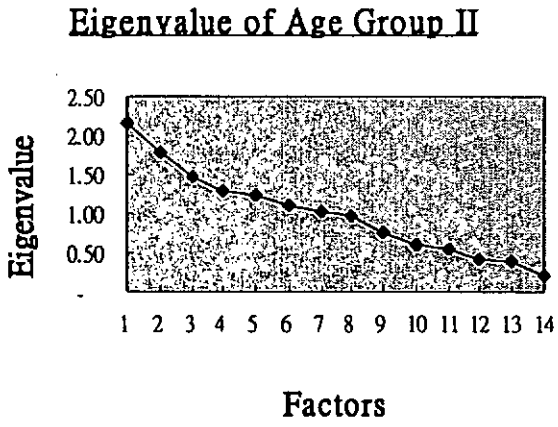
Initial Factor Analytic Statistics for CMAP for Age Group II

Items	Communality	Factors	Eigenvalue	% of variance	Cumulative %
Tower	1.00	1	2.6330	23.9	23.9
Tongue movement	1.00	2	1.5681	14.3	38.2
Follow direction	1.00	3	1.3762	12.5	50.7
Articulation	1.00	4	1.2557	11.4	62.1
General information	1.00	5	1.0044	9.1	71.2
Sentence repetition	1.00	6	0.8890	8.1	79.3
Digit repetition	1.00	7	0.8077	7.3	86.7
Finger localization	1.00	8	0.7151	6.5	93.2
Vertical writing	1.00	9	0.4457	4.1	97.2
Stepping (distance)	1.00	10	0.2859	2.6	99.8
Stepping (error)	1.00	11	0.0195	0.2	100.0

Scree plot (Figure 4.1) of successive eigenvalues against their ordinal position suggested a three factor structure according to the scree rule suggested by Nunnally and Bernstein (1994). Factor analysis was repeated to obtain the final statistics (Table 4.17).

Figure 4.1

Scree Plot of Eigenvalues (Age Group II)



The three factors selected constituted 50.7% of the total variance. The final statistics showing the communalities and the percentage of variance with three factors extracted are shown in Table 4.17. Factors 1, 2 and 3 explained 23.9%, 14.3% and 12.5% of the total variance respectively.

Table 4.17

Final Factor Analytic Statistics for Items in CMAP Sub-scales for Age Group II

Items	Communality	Factors	Eigenvalue	% of variance	Cumulative %
Tower	.42591	1	2.6330	23.9	23.9
Tongue movement	.56038	2	1.5681	14.3	38.2
Follow direction	.49312	3	1.3762	12.5	50.7
Articulation	.48320				
General information	.44237				
Sentence repetition	.77133				
Digit repetition	.75256				
Finger localization	.59084				
Vertical writing	.30164				
Stepping (distance)	.27562				
Stepping (error)	.48028				

Varimax rotation was done with the factor loadings shown in Table 4.18. Factor 1 mainly loaded with all four items of verbal sub-scale. Factor 2 consisted of two items of foundations sub-scale (finger localization and stepping (distance). Unexpectedly the tongue movement item of coordination sub-scale was loaded on this factor too. Factor 3 mainly loaded with the tower and articulation items from the coordination sub-scale. However, the stepping (deviation) item of foundations sub-scale loaded on factor 3 instead of factor 2. The vertical writing item, which was an overlapping item both in foundations and coordination sub-scale of in original MAP loaded equally in factor 2 and 3.

Table 4.18

Varimax Rotated Factor Loadings of CMAP for Age Group II

Sub-scales	Items	Factor 1	Factor 2	Factor 3
Coordination	Tower	-.30479	-.20621	.53897
	Tongue movement	-.02305	.74803	-.01747
	Articulation	.07291	.05373	.68920
	Vertical writing*	-.09027	.34742	.41568
Foundations	Finger localization	.24685	.72506	-.06473
	Stepping (distance)	-.29857	.40757	.14268
	Stepping (deviation)	.21977	.02044	.65693
Verbal	General information	.60532	.24917	.11778
	Follow direction	.68889	.03261	-.13227
	Sentence repetition	.86403	-.14929	.04992
	Digit repetition	.85781	-.08358	.09865

*item overlapping in the sub-scales of MAP

The factor loadings were very distinct, in most of the items, to their related factors. Factor 1 can be interpreted as the verbal sub-scale, whilst both factors 2 and 3 were loaded with a mixed composition of items belonging to coordination and foundations sub-scale. In considering the pattern of composition, factor 2 could be interpreted as the foundations sub-scale and factor 3 as the coordination sub-scale.

Factorial Structure of CMAP at Age Group IV

For age group IV, a total of 13 items (with two measures in stepping item: distance and deviation) were used for the final analysis. A total of 60 subjects were involved in the final analysis. The Kaiser-Meyer-Olkin measure of

sampling adequacy was 0.387 (p = 0.005) which indicated that the sample size was adequate for factor analysis. The initial statistics including the eigenvalues and the percentages of variance of the common factors are shown in Table 4.19.

Table 4.19

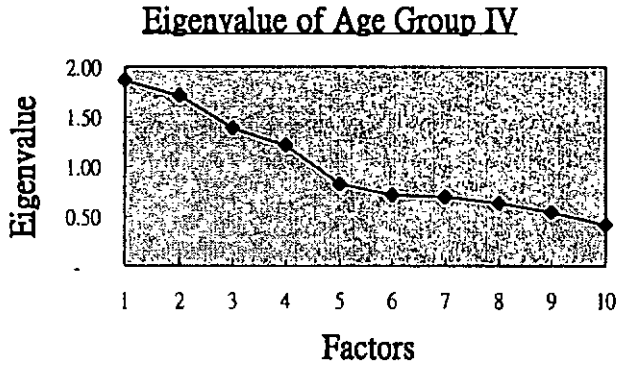
Initial Factor Analytic Statistics for CMAP for Age Group IV

Items	Communality	Factors	Eigenvalue	% of variance	Cumulative %
Supine flexion	1.00	1	2.1477	15.3	15.3
Tower	1.00	2	1.7935	12.8	28.2
Tongue movement	1.00	3	1.4799	10.6	38.7
Rapid alternate movement	1.00	4	1.2856	9.2	47.9
Follow direction	1.00	5	1.2324	8.8	56.7
Articulation	1.00	6	1.0950	7.8	64.5
General information	1.00	7	1.0392	7.4	72.0
Sentence repetition	1.00	8	0.9727	6.9	78.9
Digit repetition	1.00	9	0.7705	5.5	84.4
Finger localization	1.00	10	0.6026	4.3	88.7
Motor accuracy	1.00	11	0.5541	4.0	92.7
Vertical writing	1.00	12	.4243	3.0	95.7
Stepping (distance)	1.00	13	.3963	2.8	98.5
Stepping (deviation)	1.00	14	.2060	1.5	100.0

Scree plot of successive eigenvalues against their ordinal position was plotted (Figure 4.2). The four factor model was adopted according to the scree rule (Nunnally and Bernstein, 1994). The three factors model was not applicable as it only explained 38% of the total variance.

Figure 4.2

Scree Plot of Eigenvalues (Age Group IV)



The four factors selected constituted 47.9% of the total variance. The final statistics showing the communalities and the percentage of variance with four factors extracted are shown in Table 4.20. Factors 1, 2, 3 and 4 explained 15.3%, 28.2%, 38.7% and 47.9% of the total variance respectively.

Table 4.20

Final Factor Analytic Statistics of Items in CMAP Sub-scales for Age Group IV

Items	Communality	Factors	Eigenvalue	% of variance	Cumulative %
Supine flexion	.54261	1	2.1477	15.3	15.3
Tower	.44801	2	1.7935	12.8	28.2
Tongue movement	.42256	3	1.4799	10.6	38.7
Rapid alternate movement	.62167	4	1.2856	9.2	47.9
Follow direction	.45087				
Articulation	.33330				
General information	.60305				
Sentence repetition	.61903				
Digit repetition	.53549				
Finger localization	.35248				
Motor accuracy	.46904				
Vertical writing	.45048				
Stepping (distance)	.63030				
Stepping (deviation)	.22787				

Varimax rotation was done with the factor loading shown in Table 4.21. Factor 1 mainly loaded with two items from the verbal sub-scale (follow direction and general information) and two from the coordination sub-scale (rapid alternate movement and articulation). Factor 2 mainly consisted of three items from the foundations sub-scale (supine flexion, finger localization and

stepping (deviation). But two items from the coordination sub-scale (tower and tongue movement) also loaded on factor 2. Factor 3 mainly consisted of the sentence and digit repetition items from the verbal sub-scale and motor accuracy from the coordination sub-scale. Factor 4 mainly consisted of the vertical writing and stepping (distance) item of the foundations sub-scale. In these situations, vertical writing and rapid alternate movement items were actually overlapping items in foundations and coordination sub-scale of MAP .

Table 4.21

Varimax Rotated Factor Loadings of Items in CMAP for Age Group IV

Sub-scales	Items	Factor 1	Factor 2	Factor 3	Factor 4
Coordination	Tower	-.04466	.62420	.27073	.11580
	Motor accuracy	-.08143	.27576	.56291	.26362
	Tongue movement	-.08972	.63955	.04714	-.05707
	Articulation	.47540	-.05871	.12045	.29890
	Vertical writing*	-.07985	-.11512	.21192	.62124
Foundations	Finger localization	.34146	.38679	-.11471	.27041
	Supine flexion	.35373	.59007	-.25182	-.07677
	Rapid alternate movement*	.72757	-.19853	-.02323	-.22880
	Stepping (distance)	-.06657	-.19390	.16679	.74863
	Stepping (deviation)	.19711	.39289	.15320	.10576
Verbal	General information	.76823	.00256	.10867	.03257
	Follow direction	.44030	.34649	.36998	-.00793
	Sentence repetition	-.11927	.02323	.69305	.35206
	Digit repetition	.26957	.11750	.65999	-.11590

*item overlapping in the sub-scales of MAP

In the case of age group IV, the factor loadings in each factor were fairly clear. Two items loaded evenly across two to three factors: finger localization and follow direction. Factor 1 was a mixture of foundations and verbal sub-scales. Factor 2 was loaded with items from the coordination and foundations sub-scales. Factor 3 was mainly loaded with items from the verbal sub-scale. Factor 4 contained items from both coordination and foundations sub-scales. The two overlapping items did not load evenly across two factors.

Factorial Structure of CMAP at Age Group VI

For age group VI, a total of 11 items were used for the final analysis. The number of subjects involved was 64. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.558 (p = 0.014) which indicated that the sample size was adequate for factor analysis. The initial statistics including the eigenvalue and the percentages of variance of the common factors are shown in Table 4.22.

Table 4.22

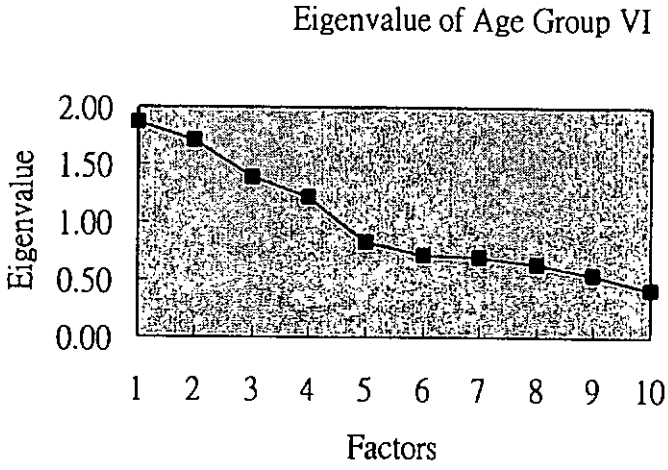
Initial Factor Analytic Statistics for CMAP in Age Group VI

Items	Communality	Factor	Eigenvalue	% of variance
Tower	1.00	1	1.8639	18.6
Follow direction	1.00	2	1.7091	17.1
Articulation	1.00	3	1.3876	13.9
General information	1.00	4	1.2146	12.1
Sentence repetition	1.00	5	0.8249	8.2
Rapid alternate movement	1.00	6	0.7137	7.1
Stereognosis	1.00	7	0.6954	7.0
Stepping (error)	1.00	8	0.6326	6.3
Digit repetition	1.00	9	0.5427	5.4
Vertical writing	1.00	10	0.4154	4.2

Scree plot (Figure 4.3) of successive eigenvalues against their ordinal positions suggested a four factor structure according to the scree rule. Factor analysis was repeated to obtain the final statistics (Table 4.23).

Figure 4.3

Scree Plot of Eigenvalues (Age Group VI)



The four factors selected constituted 61.8% of the total variance. The final statistics showing the communalities and the percentage of variance with four factors extracted are shown in Table 4.23. Factors 1, 2, 3 and 4 explained 18.6%, 17.1%, 13.9% and 12.1% of the total variance respectively.

Table 4.23

Final Factor Analytic Statistics for Items in CMAP Sub-scales for Age Group VI

Items	Communality	Factors	Eigenvalue	% of variance
Tower	.61686	1	1.8639	18.6
Follow direction	.64596	2	1.7091	17.1
Articulation	.67445	3	1.3876	13.9
General information	.62932	4	1.2146	12.1
Sentence repetition	.67456			
Rapid alternate movement	.65158			
Stereognosis	.53855			
Stepping (error)	.59150			
Digit repetition	.55520			
Vertical writing	.59729			

Varimax rotation was done and the factor loadings are shown in Table 4.24. The result at age group VI were similar to that at age group II under the verbal sub-scale. Factor 1 mainly consisted of all four items from the verbal sub-scale except the digit repetition item which also loaded on factor 2. Factor 2 mainly consisted of the two items which overlapped according to the MAP under foundations and coordination sub-scales. Factor 3 loaded with two items in the foundations sub-scale (stereognosis and stepping (deviation)). Factor 4 mainly consisted of two items under coordination sub-scale (tower and articulation).

Table 4.24

Varimax Rotated Factor Loadings of Items in CMAP for Age Group VI

Sub-scales	Items	Factor 1	Factor 2	Factor 3	Factor 4
Coordination	Tower	.16864	.16886	.05950	.74590
	Articulation	-.01418	-.12678	-.09760	.80539
	Vertical Writing*	-.06826	.76673	-.03394	.06002
Foundations	Stereognosis	.06926	-.12679	.71938	.01302
	Stepping (deviation)	.03177	.02862	.76706	-.03592
Verbal	General Information	.67648	-.10153	.39347	.08105
	Follow Direction	.69067	-.07785	-.38215	.12972
	Sentence Repetition	.80660	.05838	.14205	.01913
	Digit Repetition	.36113	.52564	-.21837	-.31749

*item overlapping in the sub-scales of MAP

The factor loadings were also very distinct except digit repetition item. Factor 1 was the verbal sub-scale. Factor 2 contained only one item, vertical writing, which was an overlapping item suggested by MAP but was not overlapping with other factors in this age group. Factor 3 was the foundations sub-scale. Factor 4 was the coordination sub-scale.

In summary, items under the verbal sub-scale loaded under same factor in age group II and VI. However, for age group IV, where some limitations arised during designing of the new items, the items loaded on two factors evenly. In all three age groups, the items under the coordination sub-scale also tended to cluster together under one or two sub-scales. The items under the foundations sub-scale scattered across the factors more and the factor loadings were not very discrete.

CHAPTER V

DISCUSSION

Introduction

This chapter discusses the principles in translating language aspect of assessment for preschool children with Cantonese as their spoken language. Particular examples will be drawn from the experience, results of the panel reviews and findings from the field tests. The standardization of three CMAP sub-scales with consideration of the cultural differences between United States and Hong Kong children will be further evaluated. The psychometric properties of the three CMAP sub-scales in terms of its content related, structural, and construct validity will be presented. Different norm tables and error estimation of the CMAP for age groups will also be presented for reference purpose.

Test Translation from MAP to CMAP

This study involved translation of MAP into its Cantonese version, CMAP. This translation process enabled the English based MAP to be applied in the screening of developmental delay children in Hong Kong. Hong Kong is a country with over 90% of the citizens are Chinese (Hong Kong Government, 1996). The spoken language used is Cantonese which is one of the most widely known "yue" group of dialects. "Yue" dialects are spoken primarily in southern Chinese provinces of Guangdong and Guangxi, including Hong Kong and Macau. Varieties of Cantonese are also used in other Chinese communities in Singapore, Malaysia, North America, Australia and elsewhere

as a result of emigration from the Guangdong area (Matthews & Yip, 1994). Although people in Hong Kong speak in Cantonese, the standard written language is primarily based on Putonghua in terms of grammatical structure. There is a close relationship between Putonghua and Cantonese in terms of grammatical structures. However, the differences in phonology, vocabulary and sentence structure between the two languages are subtle. For example, in sentence structure, the noun representing the agent of the action must be present in Cantonese while in Potunghua, the agent can be omitted (Matthews & Yip, 1994). Due to these subtle differences, translation of the instrument to Putonghua implies that the instructions are translated into the written language of Hong Kong people. Although testers in Hong Kong can read the instructions, they have to convert them into their spoken language for the preschool children. During this conversion process, testers may used different levels of Cantonese vocabularies and grammatical structures in which the difficulty level different for different age groups. Therefore, it is essential to translate MAP into CMAP and establish the equivalence between the two versions.

Language Structure of Cantonese

Phonology of Cantonese

Phonology is the study of the sounds of language. In Cantonese, the phoneme in a single word which can be divided into "initial-vowel-finals". Cantonese is different from English, a different word can be made by changing the tone in the word (e.g. yau; "worry" in English and yauh; "again" in English) or the different length of the vowels (e.g. san "new" in English versus saan "hill" in English) (Matthews & Yip, 1994). Altogether there are nineteen initials (onsets) and eight finals (codas) (So, 1993). An example of the onsets is "m"

as in "ma" (mother in English) and an example of the codas is "-t" as in "baat" (eight in English). For detail breakdown of the onsets and their related words, please refer to articulation item of Appendix F. According to a study by So (1993), all these phonemes are acquired by children in Hong Kong before the age of six (So, 1993). Therefore, these phonemes can be used to assess children's phonology ability before age of six.

Variation and Change in Pronunciation - Among these nineteen onsets, Matthews and Yip (1994) concluded that the pronounciation "n" (for example: "neih" as "you" in English) and "ŋ" (for example: "ngoh as "I" in English are currently less commonly used by the younger generation. These onsets are found to be replaced by the onset "l" as in "leih" instead of "neih" and zero onset as in "oh" of "ngoh".

Having these in mind, in stage one of the preliminary field test, over 75% of the children were found to present problems in these two onsets. In the articulation item, children were required to repeat the words said to them, this requires the children to compare the auditory information received with the information in the output lexicon before they can reproduce it (Ng, 1995). Thus, the children who failed to reproduce the "n" and "ŋ" onsets may not have such onsets stored in their output lexicon. It is therefore suggested that extra caution should be taken when interpreting children who failed to reproduce these two onsets. In addition, in spoken language, such as Cantonese, minor variation and change in pronunciation may be present with change in generations (Bourgerie, 1990). Matthews and Yip (1994) also observed that the onsets of "gw" as in "gwo" (piece in English) and "kw" as in "kwo"(country in English) seemed to be changing to "g" as in "go" and "k" as in "ko". Although

the children in this study did not demonstrate difficulty in pronouncing these two onsets. With these variations and changes, one should take extra caution.

Choice of onsets or codas - In the MAP, the articulation item requires the children to accurately reproduce the initials and the finals of the real words. For example: "mom" and "good" assessed four phonemes: initals: "m" and "g" and finals: "g" and "d". However, in Cantonese, some of the finals are unreleased (Matthews and Yip, 1994). For example, in "faat - law", the -t is formed by the tongue touching the alveolar ridge behind the teeth, but without air being released. The codas: -p, -t and -k, which are all unreleased consonants, are very difficult to distinguish for therapists not specialized in speech. In this research, both the researcher and the research assistants had difficulties in assessing these finals (codas) which are part of the articulation items. The eight codas were included in preliminary field test stage one (Appendix F). However, the inter-rater reliability among the researcher and three research assistants were very low, especially in these unreleased consonants. On the other hand, by using two words in assessing each onset, the total number of phonemes (34 phonemes) used simulated that of MAP (25 - 36 phonemes). Thus, only the onsets are used in CMAP.

Grammatical Structure of Cantonese

Cantonese grammatical structure was largely investigated by Cheung (1972), Ko (1980), Ko et al (1985) and Matthew and Yip (1994). Among them, Metthews and Yip (1994) had made a very detail analysis on the structure of Cantonese. But they did not specify the age of children in acquiring these grammatical structures. Ko et al (1985), while developing the Chinese Reynell Developmental Language Scale (CRDL) for children in Hong Kong, had

compared the difference in the structure and vocabulary of Cantonese and English based on the test content of the CRDL. Supplement to the analysis of Matthews and Yip (1994), they identified the developmental level of some of the language performances of preschool children in Hong Kong. In the process of translating and designing new questions for the language specific test items in MAP and the findings from the preliminary field test, some additional findings to the developmental level of language performance can be made.

Difficulty Level in Different Language Structures

In writing questions for CMAP sub-scales, the difficulty level of each question was often difficult to estimate as Cantonese is a spoken language, and many of its grammatical structures had not been systematically verified (Matthew & Yip, 1994). This makes it difficult for the test developer and translator to translate questions with equivalent difficulty levels as their English counterparts. In designing questions for the preschoolers, the situation was further complicated by the developmental level of the children being different for different age ranges. Though CMAP was not a language test per se, in the process of translating its verbal sub-scale, the difficulty levels in the instructions while using localizers, adjectives, and vocabularies, needed to be identified.

Localizers/Postpositions - Localizers are expressioins of location in Cantonese. In English grammar, the prepositions, which are words used with a noun or pronoun to show its connection with another, include the function of the localizers. Localizers serve two main functions: as adverbs of location and following a noun phrase, used in the conjunction with the locative coverb "hai"

("at" in English) (Matthews & Yip, 1994). In this research, the localizers of "side-by-side" and "inside" are easy for children at age group II (item difficulty at 0.86 and 0.82 respectively). For age group IV, they have generally acquired the concept of "back" (item difficulty at 0.67). For age group VI, they have the concept of "centre of the object" (item difficulty at 0.92).

Vocabularies - Vocabularies refer to the children's ability to tell the meaning of a specific word accurately. During this research, some vocabularies in each age group were identified and there were age differences in performances shown. For age group II, their acquired vocabulary included bed and towel (item difficulty at 0.90 and 0.88 respectively). For age group IV, their vocabularies included ear and firemen (item difficulty at 0.93 and 0.95) . For age group VI, they were able to explain the meaning of washing hands in at least two steps (item difficulty at 0.89). The difference in the original MAP lies in the word: firemen. Children acquired this word at age group IV instead of age group VI. With reference to their respective books at kindergarten, the picture of fireman was commonly found. Moreover, with such a high density of population in Hong Kong, it is very common to see a fire engine passing by or fireman on television. Thus, Hong Kong children may be more familiar with such an occupation.

Classifiers - Classifiers are used to classify nouns by features such as shape, size, and function. They are used in counting, like four pairs of shoes in English, but are used much more extensively than their English counterparts and are required in many contexts. For example: saam jek gau (three CL dog) as three dogs in English (Matthews & Yip, 1994). In CMAP, four classifiers are used "go", "jek", "kin", "gi" and "ka" in different age groups. However, as they

were assessed with the whole sentence, no specific age difference can be identified.

Perfectives - In Cantonese or Chinese, there is no specific tense structure. Instead, it uses perfectives to express the notion of a complete event (Matthews & Yip, 1994). In CMAP, perfectives were used in sentence repetition item. There is no obvious difference in performances across age groups.

In general, the difficulty level increases with an increase in the number of characters and steps involved and number of concepts such as the vocabulary, localizers, and perfectives in the sentences. These findings give supplementary information in designing or translating developmental assessment instruments for children in Hong Kong in follow direction, sentence repetition and general understanding in concepts and vocabularies.

Translation of Instructions for Preschoolers in Hong Kong

During the translation process, which included translation, backward translation, and test of equivalence by expert panels, there were some controversial points for further discussion. Backward translation had been recommended by Satorius and Kuyken (1994) as one of the steps during translation of an instrument. However, Geisinger (1994) and Hambleton (1993) had a different view point on this procedure. They believed that backward translation precluded the use of meaningful and appropriate substitutions of item content for cultural reasons. In this study, backward translation and test of equivalence were conducted to maximize the detainment of meaningfulness and appropriateness as compatible to the original MAP. Despite the fact that backward translation was able to ensure exact translation, language

appropriateness and grammatical structure and understandability of the CMAP solely relying on direct translation were problematic. Many differences in the grammatical structure of English and Cantonese affected the understandability and fluency of the translated version. Test of equivalence between the source (English) and target (Cantonese) versions were found to be more useful. In this process, over 70% of the items needed some alterations in their instructions. The understandability of the translated instructions was then much improved.

Special precaution should be taken in the number of members in an expert panel. In this study, it was not uncommon to encounter panel members presenting very different opinions. This is usually because people have different habits in a spoken language, as Cantonese. For example, in vertical writing item, four members suggested the word "遞" (dai) while two members suggested "吊" (diu) to replace the English word: "keep your arm up". The final decision was made with only one members' difference in opinion. The number of the members in the expert panel may increase so as to reach consensus with a good percentage of agreement of over 70% required. Nevertheless, the panel size should still allow discussions to take place (Geisinger, 1994).

Psychometric Properties of CMAP

Standardization of CMAP

Standardization means uniformity of procedures in administering and scoring the test. In order to secure uniformity of testing condition, test developer is required to provide detailed instructions for administering the test. Such standardization extends to the exact testing materials, time limits, verbal

instructions, preliminary demonstrations, response from test takers, and descriptions of testing situation (Anastasi & Urbina, 1997). In the context of test translation, each of these components of the CMAP will be evaluated. Furthermore, children's test taking behaviour attributable to cultural differences between United States and Hong Kong children, as reflected in this study, will be addressed.

Administration Procedures of CMAP

Except the verbal instructions and test content of the four language specific items, all the directions for administrating CMAP followed the guidelines stipulated in the original CMAP. These directions include instructions, demonstrations and clarifications needed in case of doubt. In the pilot test of this study, the clarity was reviewed by the researcher and three research assistants. Children participating were asked to comment on their understandability during the pilot test. Responses from the children and the comments from the research assistants were positive. They found that the directions were very clear and user friendly. No additional instructions or directions in test administration were therefore required.

Equipment and Materials of CMAP

Two pencils with different lengths and one cup were added to the original MAP in order to carry out the follow direction item. For example: at age group VI, a new question "Put the shortest pencil beside the cup" was used to replace the old question "Turn over the penny. Give me the pencil. Tell me your name." All the equipment and materials are included in the MAP assessment kit and were found to be relevant for use in testing the children in Hong Kong.

Administrative Time

The original MAP suggested the test completion time to be 25 - 35 minutes. However, in this study, about 30 minutes was required to complete 18 items of three sub-scales. For a complete test of CMAP, which contains 27 items of five sub-scales, it is predicted that the time of administrating for CMAP should be at least 45 minutes for children in Hong Kong. This is regarded as rather long for administering a screening test in clinical situations. As proposed by Aylward (1994), a screening test should take 15 - 20 minutes to complete as it is usually administered to a large number of normal children as a preassessment. Hence, if CMAP is used as a screening instrument, it has to be shortened.

Content-related Validity of CMAP

Results from two expert panel reviews (six members each) suggested that CMAP was relevant and representative to assess children with developmental delay in Hong Kong. However, there were two members (33%) who did comment on the difficulty level of two items: hand-to-nose and kneel-stand being too easy to assess children in Hong Kong. The comments had been compromised and new items were not prepared before the final field test. As reviewed in the field test, these two items reached the ceiling and needed to be revised. It is therefore suggested that additional items should be made with any members' comment and another expert panel meeting to be held to comment on that additional items before actual commencement of the field test (Smith & McCarthy, 1995).

Construct Validity of CMAP

Test Taking Behaviour of Hong Kong Children

In the final field test of this study, interesting phenomenon was observed which reflected differences in the test taking behaviour of children in Hong Kong and United States. Test taking behaviour is believed to influence the performance on a test (Anastasi & Urbina, 1997) and hence scores on a test will be different. The percentile scores of the normal children participated in the protocol of CMAP and those in the United States using MAP were compared. This is because the raw scores of United States children is not available. The scorings in each item were converted to standard scores from 0.00 to 1.00 for the ease of comparison. As follow direction, general information, articulation and sentence repetition items are newly designed items, they are deleted for comparison. In comparing the percentile scores of children between two places, Hong Kong children performed the same (8%) or better (58%) in the 64 scorings (Table 5.1). In the comparison, it was found that under the same walks line item, Hong Kong children performed poorer in terms of speed during the walks line assessment but committed less errors.

Table 5.1

Comparison of Test Performance by Children between Hong Kong and United States

Items	Standard score (0.00 - 1.00) US Scores (HK Scores)					
	Age Groups					
	II		IV		VI	
	5th % tile	25th % tile	5th % tile	25th % tile	5th % tile	25th % tile
Tower	.38 (.44)	.50 (.63)	.44 (.44)	.63 (.56)	.50 (.52)	.75 (.69)
Stereognosis	.50 (.75)	.75 (1.00)	.50 (.75)	.75 (1.00)	.25 (.50)	.50 (.81)
Finger localization	.00 (.20)	.25 (.40)	.00 (.20)	.50 (.40)	.25 (.50)	.50 (.75)
Motor accuracy	.00 (.25)	.00 (.39)	.00 (.07)	.19 (.28)	.10 (.33)	.25 (.42)
Vertical writing	.00 (.53)	.45 (.71)	.65 (.22)	.35 (.59)	.55 (.25)	.35 (.63)
Hand-to-nose	.00 (1.00)	.50 (1.00)	.33 (.37)	.83 (1.00)	.50 (1.00)	.83 (1.00)
Romberg	.00 (1.00)	.93 (1.00)	.40 (.94)	.93 (1.00)	.00 (.37)	.20 (1.00)
Stepping (distance)	.00 (.75)	.75 (1.00)	1.00 (.53)	.50 (1.00)	1.00 (.40)	.50 (.80)
Stepping (deviation)	.00 (.41)	.80 (.80)	1.00 (.40)	.60 (.60)	.80 (.50)	.40 (.75)
Wallks line (speed)	1.00 (.09)	.59 (.48)	1.00 (.24)	.52 (.57)	1.00 (.11)	.69 (.39)
Walks line (error)	.00 (.71)	.50 (1.00)	.00 (.18)	.29 (.83)	.00 (.20)	.30 (.60)
Supine flexion	.00 (.19)	.00 (.52)	.00 (.07)	.27 (.28)	.00 (.28)	.40 (1.00)
Kneel-stand	.00 (1.00)	.33 (1.00)	.00 (1.00)	.50 (1.00)	.00 (1.00)	.50 (1.00)
Tongue movement	.20 (.60)	.40 (.60)	.25 (.25)	.50 (.50)	.25 (.50)	.75 (.75)
Rapid alternate movement	.00 (.08)	.60 (.15)	1.00 (.15)	.73 (.35)	1.00 (.13)	.73 (.25)
Digit repetition	.00 (.60)	.33 (.60)	.33 (.20)	.50 (.30)	.33 (.60)	.50 (.80)

* Shaded areas indicate items which children in Hong Kong scored higher than those in the United States.

The differences of standard scores in these items indicate possible differences of performances on the CMAP (and MAP) between the two group of children. As mentioned in the literature review, there are cultural differences between Hong Kong and United States children behaviour and henceforth performance on the CMAP (and MAP).

American mothers regard their children more as individuals. Children are encouraged to develop through experimentation and observation of other people and events (Mcgillicuddy-De Lisis & Subramanian, 1996). On the contrary, Chinese families, rearing their children in Hong Kong, still have the trace of Chinese tradition which emphasises the importance of family. Children are regarded more as members of the family who are required to comply with to the house rules. Obedience and compliance to instructions is very important. Faults in the child are considered as shame for the family (Char, 1980). The difference in parental practice and cultural belief system may affect the children's performance during assessment. These beliefs are deep rooted and embedded in the children's belief and value system. As a consequence, children are more obedient, conform, and have a fear of committing errors.

Under this cultural context, some of the items which Chinese children scored higher can be explained by their difference in test taking behaviour. For example, in the Romberg item, children were requested to stand still with their eyes closed. Besides assessing balance ability of children, the degree of compliance to instructions is one of the factors to determine whether the children would stand still for a period of time. Furthermore, there was a tendency to avoid committing errors even in the expense of performing slower (in walks line item). For example for age group II, the standard score of Hong

Kong children in walks line (speed) item at 5th percentile was 0.09 to that the United States children was 1.00. In walks line (error), their standard score at 5th percentile was 0.71 as compared to the United States children at 0.00.

Besides behaviours originating from family education, the formal education which children receive in Hong Kong and United States contributes to the differences in children's performance. Compared to about 95% of the Hong Kong preschool children who are studying at nurseries or kindergartens (Hong Kong Government, 1996), only 20.5% of children under five are placed in child care centres in United States. Over 50% of children in United States are cared for by parents or relatives at home (Gormley, 1995). The education provided in kindergartens and nurseries in Hong Kong advocates compliance to rules and regulations (Chow, 1993). Formal structured teaching paper and pencil tasks are included in the curriculum for children at age three to four (Chan, 1993). The effect of structured education and practice was reflected in Hong Kong children who scored higher on the motor accuracy item which requires children to write precisely within a confined space.

However, at this stage, it is difficult to conclude whether the differences are attributable to children having different test taking behaviour or the construct of the test being upset with the change in the testing population. Cultural differences have always been an interesting issue to be addressed during test development and translation (Geisinger, 1994). Aiken (1997) suggested that the performance of assessment in abilities can be affected by home environment, nutrition, education experiences, urban versus rural living environment and nationality. Further studies in the cultural difference with matching of socioeconomic status, living condition, and education experiences

to rule out other causal factors may confirm whether there is a generic difference between children in Hong Kong and the United States.

Meanwhile, in application of standardized assessment instruments of other countries to children in Hong Kong, the item difficulty may differ with consideration to the specific testing behaviour of being more cautious in avoiding making mistakes and better rule compliance, leading to the possibility of better performance.

Results of Item Analysis

Item Difficulty Level

Results of item analysis indicate the difficulty levels of the hand-to-nose, Romberg, kneel-stand and stereognosis items require further adjustment. Among these four items, hand-to-nose (mean (S.D.)= 5.74 (1.05) - 5.98 (0.12), score range: 0 - 6) and kneel-stand (mean (S.D.)=1.97 (0.26) - 2.00 (0.00), score range: 0 - 2) items reached the ceiling across all three age groups. The hand-to-nose item was extracted from a neurodevelopmental screening test used to screen children with motor problem (Belcher, 1996). Kneel stand and Romberg item are the part of motor movement test (Richter & Montgomery, 1988). As screening tests, item difficulty of these two items may be set so that all normal children pass (difficulty level at 1.00) while children with motor problem failed (difficulty level at 0.00). Compared to the criteria set by the MAP which is 5% of the normal children fail on the item, the item would appear to be too easy. This opinion was also shared by some of the panel members who had the impression that the items would be too easy to test on Hong Kong children.

On the whole, the difficulty levels of hand-to-nose, kneel-stand, and

Romberg items was due to the fact that these items are highly sensitive and do not fit in the 5% base rate criteria as set out in MAP. Items with higher difficulty levels satisfying the criteria of 5% are recommended to be developed to replace the existing items.

Besides those items which have marked difference in their difficulty level. As mentioned in the testing behaviours of Hong Kong children, they performed generally better by comparing their relative percentile score as that of children in United States. Thus, it may affect their performances in stereognosis and Romberg items. It is suggested that adjustment to the item difficulty level should be made.

There are several ways to increase the difficulty level of items in CMAP. Two examples in alteration of the test content of the items are discussed in the following.

In the Romberg item, there are some differences in the test content at age group VI in the original MAP. This alteration meant the children in age group VI did not perform with a ceiling effect (mean (S.D.) = 13.50 (3.20); range: 0 - 15). The same test content may be applied to children at age group II and IV with minor adjustment in the percentile score in relation to the time in performing the test. In stereognosis item, the item difficulty level could be increased by replacing the real objects such as penny, car and rubber band used in age group II to the shapes such as circle, triangle and trapezoid used in age group VI. However, all the adjustments should undergo vigorous panel review and field testing on children representing the target population (Clark & Watson, 1995). Similar to the experience in designing new items in verbal sub-scale, two to three types of objects or questions should be arranged in the

field test so that the best one, which fitted the percentile of MAP, can be selected.

Internal Consistency of CMAP

Result of item analysis revealed moderate to high inter-item and item total correlations and internal consistency for the verbal sub-scale. In contrast, low inter-item correlations; low to moderate total item correlations, and low internal consistency indices are found for the foundations and coordination sub-scales.

High internal consistency for the verbal sub-scale (r = 0.49 to 0.79) means that the newly designed items: general information, follow direction and sentence repetition correlated well with digit repetition item of the original MAP. This also serves as the indicator that items in the verbal sub-scale are likely to measure a unidimensional verbal ability of children (Clark & Watson, 1995). However, further revision for items: follow direction (total item correlation: 0.53, $p < 0.01$) and general information (total item correlation: 0.34, $p < 0.01$) at age group IV were required. It was due to the insufficient questions designed for the field test (selecting three best out of four to five questions).

The low internal consistency indices of the foundations and coordination sub-scales suggest that questions contained in the items and hence various items in each of the sub-scales did not correlate well with each other. It further reflects that the homogeneity in foundations and coordination sub-scales of the CMAP is rather weak. It implies that the foundations and coordination sub-scales are in fact not assessing one single factor or construct.

The foundations sub-scale was designed to measure the sensory motor integrative function of children. In child development, sensori-motor stage is commonly referred as Piaget's cognitive developmental stage (Crain, 1992). It

is however not common to find literature and clinical assessments separating sensori-motor integration as a discrete clinical entity. Instead, it is regarded as part of the motor or the pre-requisites for cognitive development (Shapiro, 1996).

In the coordination sub-scale, motor accuracy and walks line (error) items needed revision due to low total item correlation in most of the age groups.

Evidence has shown that the foundations and coordination sub-scales in the CMAP at best do not measure discrete functions. Instead, the two sub-scales are combinations of assessing the cognitive and motor functions of the children. Child development as compared with other assessments, which divided children's function into discrete domains such as motor, cognitive, language and social aspects (Aylward, 1994; Brown and Elksnin, 1994; Capute & Accardo, 1996; Coker, 1989; Shapiro, 1996), the foundations and coordination sub-scales of CMAP appears to be heterogeneous.

Dimensionality of CMAP Sub-scales

In addition to internal consistency, unidimensional construct of index, each of the three sub-scales was substantiated by using explorative factor analysis technique. Results of factor analytic procedure suggested that the four items loaded satisfactorily under the verbal factor except for the age group IV data. All of the factors loadings were over 0.60. According to Clark and Watson's criteria (1994), these items are regarded as excellent items to be retained as the verbal sub-scale. This finding corresponds well with results obtained from the item analysis. For the items originally grouped under coordination and foundation sub-scales, despite about half of them yielding the high factor loadings (above 0.60) with their respective factors (sub-scales),

other items either loaded equally on two factors or loaded weakly (below 0.30) on one factor.

There are three plausible reasons to account for the weak structure within the foundations and coordination sub-scales. Firstly, as revealed previously, the differences in performance on the CMAP items of children between United States and Hong Kong causing serious distorted distribution problems in the items which would hamper on factor analysis procedures or results. Secondly, the number of subjects involved in the factor analysis being smaller in sample size. Although different authors have different opinions on the sample size for factor analysis, a relatively high sample size of 200 - 300 respondents is recommended by Comrey (1988). Thirdly, in reviewing the construct of MAP in three sub-scales, it may not be measuring one construct. In addition, in current literature, the definition of the sensory motor integrative function is not clearly stated. Thus, it is quite obvious that no positive finding can be made.

Reliability of CMAP

The inter-rater reliability of CMAP was tested to 16 subjects on four raters (researcher and three research assistants). High inter-rater reliability among four raters was found (r = 0.80 to 1.00, $p < 0.05$) except rapid alternate movement item (Table 4.8). For the rapid alternate movement item, it was repeated after additional practice among the three research assistants in the method of counting the steps. The inter-rater reliability test was repeated to seven subjects with high inter-rater reliability (r = 0.83 to 0.94, $p < 0.001$). It is therefore recommended that for training new users, emphasis should be made to the rapid alternate movement item in the way of counting number of steps.

The test-retest reliability of CMAP was not evaluated in the study.

Norm Table of CMAP

The norm table of CMAP in the mean scores, standard deviations and standard measurement of errors in the respective age groups II, IV and VI are included Table 7.1 to Table 7.3 in Appendix J. The item difficulty and total item correlation are listed in Table 7.4 to Table 7.6 in Appendix J.

Criterion-related Validity of CMAP

The original MAP was designed to identify and screen preschool children with developmental delay. At the same time, results of the assessment should reflect the performance of the children in the format of strength and weakness in five sub-scales: foundations, coordination, non-verbal, verbal and complex tasks. It is suggested that the profile can assist the therapists to develop a treatment plan (Miller, 1988).

Fotheringham (1983) suggests four levels of assessment of preschool children having developmental delay: preassessment phase; formal assessment phase one (behavioural specification); formal assessment phase two (process determinants and influences); and formal assessment phase three (recommendations for remediation). The preassessment phase aims at screening. Behavioural specification phase (formal assessment phase one) focuses on identifying the behavioural performance in respect to developmental milestones. Process determinants phase (formal assessment phase two) clarifies specific disorder and the causes of the disorder. Remedial phase (formal assessment phase three) translates the assessment findings into a usable series of recommendations for remediation of the deficits that have been identified. These recommendations are reached based on the

findings from assessment phases one and two. According to the aims specific in the CMAP, the assessment can be classified as preassessment and phase one assessment (behavioural specification).

CMAP as Preassessment Instrument

The CMAP targets to identify children who fall below 25th (at risk) and 5th (developmental delay) percentiles. Thus, it can be viewed as a preassessment instrument according to Fotheringham's definition (1983).

As a preassessment instrument, the sensitivity of CMAP is very important (Belcher, 1996). Sensitivity is the proportion of children who have abnormalities (true positive) and are identified appropriately (screen positive). Specificity is the proportion of children without abnormalities (true negative) with screen negative. Because of the brevity of screening tests, high sensitivity often is achieved at the cost of reduced specificity (Belcher, 1996). Besides sensitivity, the effectiveness of the screening can also be evaluated by its predictive validity. During this research, the sensitivity and predictive validity of CMAP are not tested. Thus, it is recommended further research is needed in the CMAP.

MAP as Behavioural Specified Assessment Instrument

Behavioural specification is achieved by two methods of testing: standardized measures of global or broad areas of cognitive abilities, which are used to determine the degree of deviation from population standards; and observations of performance on developmental tasks that are hierarchically ordered. Hierarchical development testing is achieved by presenting tasks within areas of development, such as fine and gross motor skills, which are arranged in the usual order in which they occur in a child's development.

These hierarchical measures are useful to program planning, primarily because they relate to tasks that caregivers want and expect children to learn (Fotheringham, 1983).

In the original MAP, items are designed under five sub-scales in hierarchical order. However, the construct of the individual sub-scale is not very clear especially in the foundations sub-scales as it loaded evenly to different sub-scales both in the original MAP (Miller, 1988) and in CMAP. The multi-faceted dimension of items under foundations sub-scale complicated the identification of specific problems of the children. In addition, the social aspect of the children, one of the major areas of children's functioning and adaptive behaviour (Shapiro, 1996), is not assessed formally in MAP or CMAP. As occupational therapists, who emphasize on the adaptation for maximal functional independence (Allen & Clark, 1990). Social adaptive function of the children facilitate therapist in designing the treatment plan. Nevertheless, items under verbal sub-scale was clearly unidimensional, which facilitated therapists to identify the language problem of the children. On the whole, its application in treatment planning may have some limitations in identifying children's functions.

It is recommended to use it as screening test in pre-assessment phase. However, further studies is required with increased sample size and adjustment to the difficulty level of individual items to reassess its structural validity in all five sub-scales. In addition, the predictive validity and the sensitivity of CMAP should be studied to ensure that it will be a good instrument for screening preschool children.

CHAPTER VI

CONCLUSION

Introduction

In the previous chapters, the results of translation and validation of 18 items of Miller Assessment for Preschoolers to its Cantonese version (CMAP) were presented. In this chapter, major findings of this study will be summarized. The clinical implications and limitations of this study will be discussed.

Designing and Translating Assessment Instruments for Children in Hong Kong

In translating or designing instruments for assessing children whose spoken language is Cantonese, several issues should be taken into consideration. First, an understanding of the language structure of Cantonese in terms of phonology and grammar is essential. The difficulty levels of the words and sentences used in the test and their equivalence to the English counterpart (the source instrument) are critical issues to be ascertained by the researcher working on the target instrument. Moreover, one has to recognize that the level of language understanding, vocabulary, use of localizers (prepositions) and perfectives differ as the child grows. Difficulty level of a sentence, either instructions or item content, increases with the length, the concepts involved and the vocabulary used. Second, the use of Cantonese words and sentences when designing or translating a test varies with time. For example, some of the phonemes which was used in tests previously designed,

say ten to twenty years ago, were found to become unpopular to the youngsters. In Cantonese, the "n" and "ŋ" onsets are not used commonly in young people now. Third, although backward translation can ensure equivalent meaning of words/sentences to be maintained, when the source instrument was translated to the target instrument (such as MAP to CMAP), this process was found to be not adequate for establishing equivalence between the two test versions. Expert panel reviews demonstrated a very important function of evaluating the semantic meaning and fluency of the translated version. Panel review is also unique in evaluating the grammatical structure and difficulty level associated with the language used by the target population.

Developmental Performance of Hong Kong Children

Findings in the study showed that there were differences in children performance on MAP and CMAP between Hong Kong and United States. Hong Kong children tended to act very cautiously during the walks line and Romberg items. This difference may be attributable to the influence and expectation of the children's parents between the two places. Parents in Hong Kong were found to expect their children to be obedience and compliant to the instructions provided to them. On the other hand, parents in United States expect children to be explorative and experimental in the testing environment.

In addition to parental expectations, early education provided in nurseries and kindergartens may have an effect in the resulting test taking behaviour of children in Hong Kong. In Hong Kong, children generally commence schooling at an earlier age of three. Whilst in the United States, children commence

schooling at the age of five. The familiarity to paper and pencil tasks in schools may enhance children's performance on the motor accuracy item in CMAP.

Thus, the difficulty level of the test should be adjusted correspondingly by replacing items with higher difficulty levels and cultural relevance in the process of translating an assessment originated from United States into Cantonese. Special attention should be put to translate items which involve paper and pencil tasks and tasks which tap on compliance such as accuracy for children in Hong Kong.

Psychometric Summaries on Properties of CMAP

In studying the psychometric properties of CMAP, high inter-rater reliability was found in all the 18 items under the three sub-scales. Evaluation by expert panels revealed that all the items were relevant to and representative of the assessment of Chinese children with developmental delay in Hong Kong. However, difficulty levels of some of the items require further adjustment in deleting or replacing: hand-to-nose, Romberg, kneel-stand and stereognosis. The test construct of the verbal sub-scale was demonstrated as unidimensional whilst the test construct of foundations and coordination sub-scales were not clearly differentiated. Further analysis of the results obtained revealed the motor, sensory and coordination functions assessed by the CMAP were heterogeneous to one another. Hence, the profile as stipulated by the original MAP (and also CMAP), in terms of five sub-scales, may not be as informative as what it should be. Clinicians should take extra caution when interpret the percentile and profile scores. Instead, clinicians and test users are suggested to also observe children's performance

during the test. Clinicians should interpret the CMAP test results with the consideration of children's actual behaviour in the test and test taking behaviour.

With the use of cluster random sampling of 184 children in this study, the normative data of CMAP represents the performance of preschool children in the age ranges: three years and three months to three years and eight months (age group II); four years and three months to four years and eight months (age group IV); and five years and three months to five years and eight months (age group VI). The normative data can be used as a baseline for developing further studies on the application of the CMAP. For example, testing the sensitivity of CMAP as a screening instrument to identify children with developmental delay; development of a shortened or brief version of CMAP to reduce the administration time; and studies to test the predictive power of CMAP.

Application of CMAP to Local Practice

As the non-verbal and complex tasks sub-scales were translated and validated previously, the present study on validating the foundations, coordination, and verbal sub-scales thus completed the translation of the MAP into CMAP. Normative data based on cluster random sampling of 184 children are available for the foundations, coordination and verbal sub-scales at age groups II, IV and VI.

Consistently high inter-rater reliability were founded in all sub-scales except extra practice time may be required for therapists who are with less experience in counting number of alternate steps in the rapid alternate

movement item. The test contents of the verbal sub-scale and the articulation item under the coordination sub-scale were adapted for the children in Hong Kong whose mother tongue is Cantonese. Its dimensionality and construct were confirmed by item analysis and factor analysis with very satisfactory results. However, the general information and follow direction items for age group IV require further revision in terms of its difficulty level.

In terms of difficulty level of the items, there was a general trend of the children in Hong Kong to perform consistently better in table tasks. Thus, there was a shift of higher standard scores at the 5^{th} and 25^{th} percentile for children in Hong Kong than those in the original MAP. In particular, the hand-to-nose and kneel-stand items were found to reach their ceiling scores for all the three age groups suggesting the needs to consider either to eliminate the items or replace them with more difficult items in future studies for clinical interpretation. Similarly, the stereognosis and Romberg items for age groups II and IV also required further adjustment on their difficulty levels. As a consequence, clinicians are recommended to interpret the results of these four items with extra cautions.

In addition, results from the factor and item analyses revealed that the constructs of the foundations and coordination sub-scales appeared to overlap across two latent factors. This suggested that items grouped under these two sub-scales may not form a single construct. In fact, the original MAP also revealed similar findings. Hence, it is suggested that clinicians should not single out the foundations and coordination sub-scales as unique factors. The items under the foundations sub-scale may reflect children's motor performance. In the same token, the items under the coordination sub-scale

are actually consist of three specific motor components: oral-motor coordination, fine motor coordination and gross motor coordination. Thus, clinicians should examine the performance of the child in their respective specific motor components at the item level before treatment planning can be made based on the assessment results.

Limitations of the Study

Due to the limited man power, in this study, results obtained would be limited to children in three age groups instead of the six age groups as in the original MAP. Age group I (two years nine months to three years two months); age group III (three years nine months to four years two months); and age group V (four years nine months to five years two months) were not included in this study. However, the difficulty level of individual items in their respective age groups of this study may provide a general guideline in designing or translating questions for children in those age groups (I, III and V).

In addition, some items, which were found to have high item difficulty or distorted distribution, were not replaced nor modified in this study. These items may affect the inter-item correlations and internal consistency within their respective sub-scales. Additional expert panel reviews and field tests should be added to delineate items with distorted distribution or item difficulty before the administration of CMAP in the final field test.

Although there were a total of 184 children recruited in this study, as they were divided into three age groups for conducting the factor analysis, the examination of the test structure of CMAP was only based on about sixty children. This is a relatively small sample size with respect to the usual

practice of at least ten subjects for each item (Floyd & Widaman, 1995). The limited sample size may affect the quality and interpretability of the results of the factor analysis.

Lastly, although previous research were conducted on the non-verbal and complex tasks sub-scales, normative data has not been established for these two sub-scales. The CMAP in this study should actually be referred to as the Cantonese Miller Assessment for Preschoolers - Revised. Further studies on the these two sub-scales would be able to add valuable information to the test and thus complete the validation of the Cantonese Miller Assessment for Preschoolers.

References

Allen, A. S., & Pratt, P. N. (1989). Occupational Therapy in Children. Missouri: Mosby

American Occupational Therapy Association. (1973). Consultation in the Community: Occupational therapy in child health. Dubuque, IA: Kendall/Hunt.

Anastasi, A. (1988). Psychological Testing (6th ed.). New York: Macmillan.

Anastasi, A. (1995). Psychology evolving: Linkages, hierarchies, and dimensions. In F. Kessel (Ed.). Psychology, Science and Human Affairs: Essays in honor of William Bevan. Boulder, CO: Westview Press.

Anastasi, A., & Urbina, S. (1997). Psychological Testing (7th ed.). New Jearsey: Prentice-Hall International , Inc.

Aylward, G. P. (1994). Practitioner's Guide to Developmental and Psychological Testing. New York: Plenum Publishing Corporation.

Ayres, A. J. (1963). Eleanor Clark Slagle Lecture-The Development of Perceptual-motor Abilities: A theoretical basis for treatment of dysfunction. American Journal of Occupational Therapy, 17, 221-225.

Ayres, A. J. (1972). Southern California Sensory Integration Tests. Los Angeles: Western Psychological Services.

Ayres, A. J. (1972). Sensory Integration and Learning Disorders. Los Angeles: Western Psychological Services.

Banus, B. J. (1983). The Miller Assessment for Preschoolers (MAP): An Introduction and Review. American Journal of Occupational Therapy, 37, (5), 333-341.

Batshaw, M. L., & Perret, Y. M. (1992). Children with Disabilities (3rd ed.). Baltimore: Paul H. Brookes Publishing Co., Inc.

Belcher, H. M. E. (1996). Developmental Screening. In A. J. Capute and P. J. Accardo (Eds.). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume II. Baltimore: Paul H. Brookes Publishing Co., Inc.

Bonavia, J., & Sterry, P. (1994). Essential Hong Kong. Chicago: Automobile Association.

Bourgerie, D. (1990). A Quantitative Study of Sociolinguistic Variation in Cantonese. Unpublished PhD Dissertation. Ohio: Ohio State University.

Brown, F. R., & Elksnin, N. (Eds.). (1994). An Introduction to Developmental Disabilities: A Neurodevelopmental Perspective. California: Singular Publishing Group, Inc.

Bruininks, R. (1978). Bruininks-Oseretsky Test of Motor Proficiency. Circle Pines, MN: American Guidance Service.

Burgemeister, B., Blum, L., & Lorge, I. (1954). Columbia Mental Maturity Scale. New York: The Psychological Corporation.

Capute, A. J., & Accardo, P. J. (Eds.) (1996). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume I. Baltimore: Paul H. Brookes Publishing Co., Inc.

Capute, A. J., & Accardo, P. J. (Eds.) (1996). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume II. Baltimore: Paul H. Brookes Publishing Co., Inc.

Campbell, S. K. (1989). Developing Norm-Referenced Standardized Tests. New York: Haworth Press, Inc.

Chan, Y. Y. (1993). Development of Writing Skills in Preschool Children. Education Papers. Hong Kong: The Faculty of Education, University of Hong Kong.

Char, E. L. (1981). The Chinese American. In Clark, A. L. (Ed.). Culture Child Rearing. Philadelphia: F.A. Davis Company.

Cheung, F. M., Lam M.C., & Chau, T. W. (1990). Caregiving Techniques and Pre-school Children's Development in Hong Kong Families. Hong Kong: Institute of Social Studies, The Chinese University of Hong Kong.

Child and Adolescent Psychiatric Working Group, Occupational Therapy Coordinating Committee. (1994). Survey on Assessment Tools in Paediatric and Child Psychiatric Occupational Therapy Department. Hong Kong: Author.

Chow, Y. M. (1993). Kindergarten and Primary School Teachers' Expectations of School Readiness in Young Children (pp67-90). Hong Kong: University of Hong Kong, Faculty of Education.

Clark, F., Mailloux, Z., & Parham, L. D. (1985). Sensory integration and children with learning disabilities. In P. N. Clark & A. S. Allen (Eds.), Occupational Therapy for Children (2nd ed., pp.457-509). St. Louis: Mosby.

Clark, L. A., & Watson, D. (1995). Constructing Validity: Basic Issues in Objective Scale Development. Psychological Assessment, 7, No.3, 309-319

Coker, S. B. (1989). Developmental Delay and Mental Retardation. New York: PMA Publishing Corp.

Comrey, A. L. (1988). Factor-analytic Methods of Scale Development in Personality and Clinical Psychology. Journal of Applied Psychology, 56, 754-761.

Crain, W. (1992). Theories of Development, Concepts and Applications. New Jersey: Prentice Hall

Daniels, L. E., & Bressler, S. (1990). The Miller Assessment for Preschoolers: Clinical Use With Children with Developmental Delays. American Journal of Occupational Therapy, 44, 48-53.

DeGangi, G. A. (1983). A Critique of the Standardization of the Miller Assessment for Preschoolers. American Journal of Occupational Therapy, 37, 407-412.

Deloria, D. J. (1985). Review of Miller Assessment for Preschoolers. In J. V. Mitchell (Eds.). The Ninth Mental Measurements Yearbook (Vol. 2, pp. 975-976). Lincoln, NE: Buros Institute of Mental Measurements.

Dunst, C. J. (1998). Sensorimotor Development and Developmental Disabilities. In Burack, J. A., Hodapp, R. M. and Zigler, E. (Eds.). Handbook of Mental Retardation and Development. New York: Cambridge University Press.

Easterbrooks, S. (1985). Language Development. in N. H. Fallen & W. Umansky. (Eds). Young Children with Special Needs (2nd ed., pp.61-133). Columbus: Charles E. Merrill Publishing Co.

Erikson, E. H. (1963). Childhood and Society. (2nd ed.). New York: W. W. Norton & Co.

Education for Handicapped Law Report. (1986-1987). EHLR DEC.508:114, Case No. 86-039. Alexandria, VA: CRR Publishing.

Fallen, W. H. (1985). Assessment: Techniques, Processes and Issues. in N. H. Fallen & W. Umansky. (Eds.). Young Children with Special Needs (2nd ed., pp.61-133). Columbus: Charles E. Merrill Publishing Co.

Folyd, F. J., & Widaman, K. F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. Psychological Assessment, 7, 286-299.

Foster, S. L., & Cone, J. D. (1995). Validity Issues in Clinical Assessment Psychological Assessment, 7, 248-260.

Fotheringham, J. B. (1983). Mental Retardation and Developmental Delay. in K. D. Paget, & B. A. Bracken. (Eds.). The Psychoeducational Assessment of Preschool Children. Orlando: Grune & Stratton, Inc.

Geisinger, K. F. (1994). Cross-Cultural Normative Assessment: Translation and Adaptation Issues Influencing the Normative Interpretation of Assessment Instruments. Psychological Assessment, 6, 304-312.

Gesell, A., Halverson, H. M., Ilg, F. L., Castner, B. M., Ames, L. B., & Amatruda, C. S. (1940). The First Five Years of Life. New York: Harper.

Gesell, A., & Ilg, F. L. (1943). Child Development. New York: Harper & Row, Pub.

Gormley, W. T. (1995). Everybody's Children. Washington D.C. : The Brookings Institution.

Grant, J. P. (1995). The State of the World's Children 1995. New York: Oxford University Press.

Grossman, H. J. (1983). Classification in Mental Retardation. Washington D.C.: American Association on Mental Deficiency.

Hainsworth, P., & Siqueland, M. (1969). The Meeting Street School Screening Test. Providence, RI: Meeting Street School, Easter Seal Society of Rhode Island, Inc.

Hall, H. J. (1918). Arts and crafts in medicine. In Proceedings of the First Annual Meeting of the National Society for the Promotion of Occupational Therapy. Towson, MD: National Society for the Promotion of Occupational Therapy.

Hambleton, R. K., & Rogers, H. J. (1991). Advances in Criterion-references Measurement. In Educational and Psychological Testing: Theory and Applications (pp.3-43). Boston: Kluwer Academic.

Hammill, D. D. (1985). Detroit Tests of Learning Aptitude - 2. Austin: Pre-Ed.

Hammill, D. D. (1990). On defining learning disabilities: An emerging consensus. Journal of Learning Disabilities, 21, 74-78.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. Psychological Assessment, 7, 238-247.

Hong Kong Government. (1979). Education Department Annual Summary 1978-79. Hong Kong: Government Printer.

Hong Kong Government. (1994). Annual Report by Department of Health 93/94. Hong Kong: Government Printer.

Hong Kong Government. (1996). Census Report: 1996. Hong Kong: Government Printer.

Kalakian, L., & Fichstaedt, C. (1982). Developmental/adapted Physical Education. Minneapolis: Burgess Publishing Co.

King-Thomas, L., & Hacker, B. J.. (1987). A Therapist's Guide to Paediatric Assessments. Boston: Little, Brown & Co.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). Illinois Test of Psycholinguistic Abilities. Urbana: University of Illinois Press.

Ko, Y. Y., Yeung Leung, Y. M., Mak, H. L., Miu Lee, K. P., Lau, K. T., Chan Chiu, L. H., Go, B. Y., Leung, K. K., Tse, C. Y., & Lee, Y. H. (1987). Reynell Language Developmental Scale (Cantonese Version). Windsor: NFER-NELSON.

Lane, S. J., Attanasio, R., & Huselid, R. F. (1994). Prediction of Preschool Sensory and Motor Performance by 18-Month Neurologic Scores Among Children Born Prematurely. American Journal of Occupational Therapy, 48, 391-396.

Lam, M.Y. (1993). Effects of Physical Education on Gross Motor Performance of Preschool Children (pp52-66). Hong Kong: University of Hong Kong, Faculty of Education.

Lam, C. C., Yau, Y. W., & Mak, H. L. (1996). Manual of Developmental Paediatrics. Hong Kong: Icon Media Co. Ltd.

Lerner, J., Mardell-Czudnowski, C., & Goldenberg, D. (1987). Special Education for the Early Childhood Years (2nd ed.). New Jersey: Prentice-Hall, Inc.

Lemerand, P. A. (1985). Predictive Validity of the Miller Assessment for Preschoolers. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Levine, R. E. (1987). Culture: A Factor Influencing the Outcomes of Occupational Therapy. Occupational Therapy in Health Care, 4, 3-16

Mailloux, Z., Knox, S., Burke, J., & Clark, F. (1985). Pediatric dysfunction. In G. Keilhofner (Ed.). A Model of Human Occupation: Theory and Application. Baltimore: Williams & Wilkins.

Mann, M. (Ed.). (1984). Library of Nation: China. New Jersey: Time-Life Books Inc.

Matthews, S., & Yip, V. (1994). Cantonese: A Comprehensive Grammar. London: Routledge.

McGillicuddy-De Lisi, A. V., & Subramanian, S. (1996). How Do Children Develop Knowledge? Beliefs of Tanzanian and American Mothers. In S. Harkness and C. M. Super (Eds.). Parents' Cultural Belief Systems. New York: The Guildford Press.

Meisels, S. J. (1989). Can Developmental Screening Tests Identify Children who are Developmentally at Risk? Pediatrics, 83, 578-585.

Messick, S. (1993). Validity. In R. L. Linn (Ed.). Educational Measurement (2nd ed., pp13-104). Phoenix: American Council on Education and Oryx Press.

Michael, W. B. (1985). Review of Miller Assessment for Preschoolers. In J. V. Mitchell (Ed.). The Ninth Mental Measurements Yearbook (Vol. 2, pp. 976-978). Lincoln, NE: Buros Institute of Mental Measurements.

Miller, L. J. (1987). Longitudinal Validity of Miller Assessment for Preschoolers Study I. Perceptual and Motor Skills, 65, 211-217.

Miller, L. J. (1988). Longitudinal Validity of Miller Assessment for Preschoolers Study II. Perceptual and Motor Skills, 66, 811-214.

Miller, L. J. (1988). Differentiating Children with School-related Problems after Four Years Using the Miller Assessment for Preschoolers. Psychology in the Schools, 25, 10-15.

Miller, L. J. (1988). Miller Assessment for Preschoolers Manual (rev. ed.). San Antonio, TX: Psychological Corporation.

Montgomery, T. R. (1996). Neurodevelopmental Assessment of School-age Children. In Capute, A. J., & Accardo, P. J. (Eds.) (1996). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume II. Baltimore: Paul H. Brookes Publishing Co., Inc.

Ng, K. H. (1995). Phonological Processing in Children with Speech Disorders. Dissertation Paper, University of Hong Kong.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychological Testing: Principles and Applications (3rd ed.). New York: McGraw-Hill.

Opper, S. (1993). Hong Kong's Preschoolers. Education Paper Faculty of Education (pp6-17). Hong Kong: University of Hong Kong.

Opper, S. (1993). Hong Kong's Young Children: Their Preschools and Families. Hong Kong: University of Hong Kong.

Ottenbacher, K. (1991). Statement: Occupational Therapy Provision for Children With Learning Disabilities and/or Mild to Moderate Perceptual and Motor Deficits. American Journal of Occupational Therapy, 45, (12), 1069-1073.

Piaget, J. (1926). The Child's Conception of the World. Savage, MD: Littlefield, Adams and Co.

Piaget, J. (1936). The Origins of Intelligence in Children. New York: International Universities Press.

Piaget, J. (1970). Piaget's Theory. In P. H. Mussen (Ed.), Handbook of Child Psychology (4th ed.). New York: John Wiley.

Portney, L. G., & Watkins, M. P. (1993). Foundations of Clinical Research, Application to Practice. Connecticut: Appleton & Lange.

Reid, D. (1987). Occupational Therapists' Assessment Practices with Handicapped Children in Ontario. Canadian Journal of Occupational Therapy, 54, 181-188

Richter, E. W., & Montgomery, P. C. (1988). The Sensorimotor Performance Analysis. Minnesota: PDP Press.

Rodwell, S. (1991). A Visitor's Guide to Historic Hong Kong. Hong Kong: Twin Age Ltd.

Rodger, S. (1994). A Survey of Assessments used by Paediatric Occupational Therapists. Australian Occupational Therapy Journal, 41, 137-142.

Rourke, B. P., & Do Ho, J. E. D. (1994). Learning Disabilities: A Neuropyschological Perspective. California: SAGE Publication, Inc.

Rubin, L., & Crocker, A. C. (1989). Developmental Disabilities: Delivery of Medical Care for Children and Adults. Philadelphia: Lea & Febiger.

Runac, M. (1985). Motor Development. in N. H. Fallen & W. Umansky. (Eds). Young Children with Special Needs (2nd ed., pp.61-133). Columbus: Charles E. Merrill Publishing Co.

Sartorius, N., & Kuyken, W. (1994). Translation of Health Status Instruments in Orley, J. and Kuyken, W. (eds). Quality of Life Assessment in Health Care Settings. Berlin: Springer-Verlag.

Schouten, P. G. W., & Kirkpatrick, L. A. (1993). Questions and Concerns About the Miller Assessment for Preschoolers. The Occupational Therapy Journal of Research, 13, 1.

Shapiro, B. K. (1992). Normal and Abnormal Development. in M. L. Batshaw & Y. M. Perret. Children with Disabilities (3rd ed.). Baltimore: Paul H. Brookes Publishing Co., Inc.

Shapiro, B. K. (1996). Neurologically based special learning disabilities. In Capute, A. J. and Accardo, P. J. (Eds.) (1996). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume II. Baltimore: Paul H. Brookes Publishing Co., Inc.

Shapiro, B. K. (1996). Neurodevelopmental Assessment of Infants and Young Children. In Capute, A. J. and Accardo, P. J. (Eds.) (1996). Developmental Disabilities in Infancy and Childhood (2nd ed.) Volume II. Baltimore: Paul H. Brookes Publishing Co., Inc.

Shapiro, B. K., Capute, A. J., & Palmer, F. B. (1987). Primitive reflex profile: A quantitation of primitive reflexes in infancy. Developmental Medicine and Child Neurology, 26, 375-383.

Shimahara, N. K., & Sakai, A. (1995). Learning to Teach in Two Cultures, Japan and the United States. New York: Garland Publishing, Inc.

Slagle, E. C. (1921). Training aids for mental patients. Paper presented at the Fifth Annual Meeting of the National Society for the Promotion of Occupational Therapy, Baltimore.

Smith G. T., & McCarthy, E. M. (1995). Methodological Considerations in the Refinement of Clinical Assessment Instruments. Psychological Assessment, 7, 300-308.

So, K. H. (1993). Cantonese Segmental Phonology Test. Hong Kong: Bradford Publishing Co.

Tobin, J. J., Wu, Y. H., & Davidson, D. H. (1989). Preschool in Three Cultures. New York: Yale University.

Tse, S. K., & Ngau, K. (1995). Education Papers 21: Chinese Language Education for the 21st Century: A Hong Kong Perspective. Hong Kong: University of Hong Kong, Faculty of Education. 1-38.

Turner, P. H., & Hamner, T. J. (1994). Child Development and Early Education, Infancy Through Preschool. Massachusetts: Allyn and Bacon.

Umansky, W. (1985). Cognitive Development. in N. H. Fallen & W. Umansky. (Eds). Young Children with Special Needs (2nd ed., pp.61-133). Columbus: Charles E. Merrill Publishing Co.

Vernon, P. E. (1973). Intelligence and Cultural Environment. Methven & Co Ltd.

Wagner, D. A., & Stevenson, H. W. (1985). Cultural Perspectives on Child Development. San Francisco: W. H. Freeman and Company.

Wiig, E. H., & Semel, E. M. (1980). Language Assessment and Intervention for the Learning Disabled. Columbus: Charles E. Merrill.

Appendix A

Miller Assessment for Preschoolers (abilities assessed, related sub-scales and items)

Abilities Assessed	Sub-scales	Items
Sensory and Motor	Foundations	Stereognosis
		Finger localization
		Vertical writing
		Hand-to-Nose
		Romberg
		Stepping (distance/deviation)
		Walks line (speed/error)
		Supine flexion
		Kneel-Stand
		Rapid Alternate Movement
Coordination		Tower
		Motor Accuracy
		Vertical Writing*
		Walks line*(speed/error)
		Tongue Movement
		Rapid Alternate Movement*
		Articulation
Cognitive	Verbal	General information
		Follow Direction
		Sentence Repetition
		Digit Repetition
	Non-verbal	Sequencing
		Block Tapping
		Object memory
		Puzzle
		Figure-Ground
Combined	Complex Tasks	Block Designs
		Draw-a-person
		Imitation of Postures
		Maze

*represents the items which overlaps in more than one sub-scale

APPENDIX B
Invitation Letters to Expert Panel A and B for Test of Equivalence

Jan 2, 1997.

Dear

Thank you for joining the expert panel A - I on the Validation of Miller Assessment for Preschoolers (MAP). You are cordially invited to attend the meeting on 16.1.97. 2:00pm at Occupational Therapy Department, Yaumatei Child Psychiatric Center.

The panel review will last for three hours. It is divided into two parts.
Part I (two hours)
You are requested to fill in a questionnaire to comment on the translation quality of the translated instructions of MAP, Miller Assessment for Preschoolers (Cantonese Version) (CMAP). Suggestions on the modification to the translated version are welcomed.
Part II (one hour)
Discussion on the items which the panel members rated "disagree" or "strongly disagree" to confirm the finalized version among the expert panel.

Please contact me at 2384[redacted] or 7116[redacted] if there is any query concerning the meeting or the research study. Looking forward to see you. Thank you again for your support to the research.

Yours sincerely,

Magdalene Poon
Occupational Therapist

Invitation Letters to Expert Panel A and B for Content Validity

Feb 10, 1997.

Dear ,

Thank you for joining the expert panel B - I on the Validation of Miller Assessment for Preschoolers (MAP). You are cordially invited to attend the meeting on _____ , at Occupational Therapy Department, Yaumatei Child Psychiatric Center.

The panel review will last for three hours. You are requested to fill in a questionnaire to comment on the relevance and representativeness of the CMAP. Afterwards, an open discussion will be made to collect your opinions and suggestions to the relevance and representativeness; and cultural relevance in using the CMAP in screening developmental delayed children in Hong Kong.

Before attending this meeting, please try out the CMAP to five Hong Kong children.

Please contact me at 2384[redacted] or 7116[redacted]#[redacted] if there is any query concerning the panel review or the research study. Thank you again for your support to the research.

Yours sincerely,

Magdalene Poon
Occupational Therapist

Appendix C

QUESTIONNAIRE ON TEST OF EQUIVALENCE (PANEL A)

VALIDATION OF MILLER ASSESSMENT FOR PRESCHOOLERS :

EXPERT PANEL A/B

Name of Expert Panel Member: _____

Work Setting: _____

Year of Experience in Child Field: _____

Date of Review: _____

INSTRUCTIONS TO ALL PANEL MEMBERS:

1. Please read the information sheet which describes the purpose of this panel review.
2. Give your consent for participating in this expert panel by signing the consent form at page 3.
3. Please read the English version and the translated version of the MAP.
4. Then fill in the questionnaire consists of questions which guides you to evaluate the fluency and semantic meaning of MAP test items after translation of the instructions.
5. A four-point rating scale is used ranging from strongly disagree (1) to strongly agree (4). Circle the number corresponding to your evaluation for each test item. Please provide your comments, justifications, or suggestions under the open-ended portion of each item.
6. This questionnaire should not take more than one hour to complete. If there is any queries, please direct them to investigators for clarification. Your participation is much appreciated.

Part I

Information for Expert Panel

The purpose of this study is to validate three sub-scales of the Miller Assessment for Preschoolers (Cantonese Version) (CMAP). This study aims at completing the translation of the MAP into Cantonese version for three sub-scales (foundations, coordination, and verbal) and validating it in terms the construct validity, content validity and item difficulties.

As MAP's main user in Hong Kong is Occupational Therapists, 6 occupational therapists with over 4 years experiences working in child field are invited to join to form an expert panel. Panel members are involved in evaluating the equivalence of the translated instructions in terms of fluency and semantic meaning.

In this expert panel A - I, panel members are requested to evaluate the instructions of the 14 items under the foundations sub-scale and the coordination sub-scale of the CMAP. The 4 items under the verbal sub-scale will be evaluated by another expert panel B formed by speech therapists. A questionnaire is designed to guide the panel members through the evaluation. Panel members are reminded to work on the questionnaire independently. Panel members are encouraged to provide comments, justifications, and suggestions for any disagreed items under the open-ended portion of each item.

Part II

Equivalence Review Expert Panel A/B
Consent Form

Project title: Validation of Miller Assessment for Preschoolers (Cantonese Version)

Investigator: Expert panel A/B - I on Miller Assessment for Preschoolers (Cantonese Version)

This research project will collect evidence on the content validity and utility of the Miller Assessment for Preschoolers (Cantonese Version) in Hong Kong.

I agree to participate in Panel A/B - I which requires me to fill in a questionnaire which evaluates the fluency and semantic relevance of the translated CMAP test items.

This study carries no risks to me. There will be no direct benefits for me. My name will not be appeared in any documents or reports. I can refuse to answer any items in the questionnaire. All information collected in this study will be kept confidential.

I am free to withdraw my consent and stop participating at any time. I have been given the chance to ask questions. I am satisfied that all my questions have been answered. I understand that my participation in this research study is strictly voluntary, and that I may withdraw at any time. I am entitled to keep a copy of this consent for my reference. If I have any questions concerning the study I can contact Ms. Magdalene Poon at 2384[redacted].

I, _____, understand the explanation of this research study and consent to participate in the expert panel meeting.

Name of Participant: _____ Date: _____

Signature of Participant: _____ Date: _____

Signature of Witness: _____ Date: _____

Part III

Panel Review A - I

Validation of Miller Assessment for Preschoolers (Cantonese Version)
Equivalence of the instructions of the 14 test items

1. Tower
   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

   1 2 3 4
   strongly disagree disagree agree strongly agree

   Please justify your rating.

   _____
   _____
   _____

   ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

   1 2 3 4
   strongly disagree disagree agree strongly agree

   Please justify your rating.

   _____
   _____
   _____

2. Stereognosis
   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

   1 2 3 4
   strongly disagree disagree agree strongly agree

   Please justify your rating.

   _____
   _____
   _____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

3. Finger Localization

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

4. Motor Accuracy

   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

   ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

5. Vertical Writing

   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

6. Hand-to-Nose

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

7. Romberg

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

8. Stepping

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

9. Walks Line

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

10. Supine Flexion

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

11. Kneel-Stand

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

12. Tongue Movement

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

13. Rapid Alternating Movement

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

14. Digit Repetition

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

1	2	3	4
strongly disagree	disagree	agree	strongly agree

Please justify your rating.

_____

_____

_____

APPENDIX D
QUESTIONNAIRE ON TEST OF EQUIVALENCE (PANEL B)
Part III

Panel Review B - II

Validation of Miller Assessment for Preschoolers (Cantonese Version), CMAP
Equivalence of the Translated Instructions, four language specific test items

1. Articulation

   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

2. General Information

   i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

3. Follow Direction

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

4. Sentence Repetition

i) Are the words used in the translated version, CMAP, presented fluently and correctly as in the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

ii) Do the words used in the translated version, CMAP, have the same semantic meaning compared with the original version?

strongly disagree	disagree	agree	strongly agree
1	2	3	4

Please justify your opinion.

_____

_____

APPENDIX E

QUESTIONNAIRE ON CONTENT VALIDITY

VALIDATION OF MILLER ASSESSMENT FOR PRESCHOOLERS

(CANTONESE VERSION)

EXPERT PANEL A/B

Name of Expert Panel Member: _____

Work Setting: _____

Year of Experience: _____

Date of Review: _____

INSTRUCTIONS TO ALL PANEL MEMBERS:

1. Please read the information sheet which describes the purpose of this panel review.
2. Give your consent for participating in this expert panel by signing the consent form at page 3.
3. Please read the English version and the translated version of the MAP.
4. Then fill in the questionnaire consists of questions which guides you to evaluate the relevance and representativeness of MAP test items in assessing children's developmental delay in Hong Kong.
5. A four-point rating scale is used ranging from poor (1) to excellent (4). Circle the number corresponding to your evaluation for each test item. Please provide your comments, justifications, or suggestions under the open-ended portion of each item.
6. This questionnaire should not take more than one hour to complete. If there is any queries, please direct them to investigators for clarification. Your participation is much appreciated.

Part I

Information for Expert Panel

The purpose of this study is to validate three sub-scales of Miller Assessment for Preschoolers (Cantonese Version) (CMAP). This study aims at validating it in terms the construct validity, content validity and item difficulties.

The Miller Assessment for Preschoolers is designed by an Occupational Therapist, Ms. Lucy Jane Miller in 1982. The aim of the assessment is to screen preschool children with mild to moderate developmental delay to predict 'preacademic' problems. The age range of the test is from 2 years and 9 months to 5 years 8 months, dividing into 6 age groups:

- Age Group I 2 years 9 months to 3 years 2 months
- Age Group II 3 years 3 months to 3 years 8 months
- Age Group III 3 years 9 months to 4 years 2 months
- Age Group IV 4 years 3 months to 4 years 8 months
- Age Group V 4 years 9 months to 5 years 2 months
- Age Group VI 5 years 3 months to 5 years 8 months.

The test contains 5 sub-scales : foundations, coordination, verbal, non-verbal and complex tasks.

In this expert panel III, panel members are requested to evaluate the instructions and content of the 10 items under the foundations sub-scale and 7 items under the coordination sub-scale of the CMAP. A questionnaire is designed to guide the panel members through the evaluation. Panel members are reminded to work on the questionnaire independently. Panel members are encouraged to provide comments, justifications, and suggestions for any disagreed items **directly** on the translated/rewrite version. Further comments can also be made to the space provided at the end of each open-ended question.

Key words in the questionnaire

"***relevance***" refers to the exploration of the conceptual equivalence of the translated version, i.e. whether the translated version is measuring the same concept as the original one. Special consideration should be made in the cultural relevance of the test item after the translation.

"***representativeness***" refers to the whether the test really assesses what it claims to measure.

Part II

Content Validity Review Expert Panel A/B
Consent Form

Project title: Validation of MAP (Cantonese version)

Investigator: Expert panel A/B - II on CMAP

This research project will collect evidence on the content validity and utility of the Miller Assessment for Preschoolers (Cantonese Version) (CMAP) in Hong Kong.

I agree to participate in Expert Panel A/B which requires me to fill in a questionnaire which evaluates the relevance and representativeness of the translated CMAP test items.

This study carries no risks to me. There will be no direct benefits for me. My name will not be appeared in any documents or reports. I can refuse to answer any items in the questionnaire. All information collected in this study will be kept confidential.

I am free to withdraw my consent and stop participating at any time. I have been given the chance to ask questions. I am satisfied that all my questions have been answered. I understand that my participation in this research study is strictly voluntary, and that I may withdraw at any time. I am entitled to keep a copy of this consent for my reference. If I have any questions concerning the study I can contact Ms. Magdalene Poon at 2384[redacted] or 7116[redacted] call [redacted].

I, _____, understand the explanation of this research study and consent to participate in the expert panel meeting.

Name of Participant: _____ Date: _____

Signature of Participant: _____ Date: _____

Signature of Witness: _____ Date: _____

Part III

Panel Review A - I

Evaluation of Content Validity of Miller Assessment for Preschoolers (Cantonese Version) (CMAP) (Foundations and Coordination Sub-scales)

Foundations sub-scale

1. Stereognosis

   How well is this item relevant to the assessment of the foundations sub-scale of the developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

2. Finger Localization

   How well is this item relevant to the assessment of the foundations sub-scale of the developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

3. Vertical Writing

   How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

4. Hand-to-Nose

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

5. Romberg

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

6. Stepping

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

7. Walks Line

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

8. Supine Flexion

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

9. Kneel-Stand

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

10. Rapid Alternate Movement

How well is this item relevant to the assessment of the foundations sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

11. How well do these 10 test items (1-10) completely represents the assessment of foundations ability of developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

Coordination sub-scale

1. Tower

   How well is this item relevant to the assessment of the coordination sub-scale of the developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

2. Motor Accuracy

   How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

3. Vertical Writing

   How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

4. Walks Line

How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

5. Tongue Movement

How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

6. Rapid Alternate Movement

How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

7. Articulation

How well is this item relevant to the assessment of the coordination sub-scale of the developmental delay children in Hong Kong

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

8. How well do these 7 test items (1-7) completely represents the assessment of the coordination ability of developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

Part III

Panel Review B - I

Evaluation of Content Validity of Miller Assessment for Preschoolers, CMAP (Verbal Sub-scale)

1. General Information

   How well is this item relevant to the assessment of the verbal sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

2. Follow Direction

   How well is this item relevant to the assessment of the verbal sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

   Please justify your opinion.

   _____

   _____

3. Sentence Repetition

How well is this item relevant to the assessment of the verbal sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

4. Digit Repetition

How well is this item relevant to the assessment of the verbal sub-scale of the developmental delay children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

5. How well do these 4 test items (1-4) completely represents the assessment of the verbal ability of developmental delayed children in Hong Kong?

Poor	fair	good	excellent
1	2	3	4

Please justify your opinion.

_____

_____

APPENDIX F

Preliminary Field Test on the Four Language Specific Items (Stage One)

1. Follow Directions:
我地而家玩一個「我講你做」遊戲。
你聽清楚我叫你做乜，你就照做，
「拍手掌」(Prompt 你識唔識拍手掌呀？或示範拍手) 好、拍得好。
我會一次過叫你做幾樣嘢。
等我講完之後，先至輪到你喎。

Age Group II
物件： 五個硬幣、一個盒、三枝筆 (不同長短)、一個杯、一個波

Q1. 擺三個銀係盒裏便。	
Q2. 將枝最長嘅筆放係盒下便。	
Q3. 係杯度拎兩枝筆出嚟。	
Q4. 係枱度搵枝筆比我。	
Q5. 拋起個波。	
Q6. 擺晒 D 銀比我。	

Age Group IV
物件： 五個硬幣、一個盒、三枝筆(不同長短)、一個杯、一架車〔頭向小朋友〕

Q1. 將全部筆擺係杯後面。	
Q2. 將兩個銀擺埋一齊。	
Q3. 擺晒 D 筆同埋銀係盒裏面。	
Q4. 擘開個口， Lam 吓你個口唇，然後笑一笑。	
Q5. 將個銀放係車仔後面。	
Q6. 掂吓你隻耳仔同埋鼻，然後繞住你張櫈行一個圈。	

Age Group VI

物件： 五個硬幣、一個盒、三枝筆(不同長短)、一個杯、一個波、一張紙

Q1. 擺短果枝筆喺杯側邊。	
Q2. 擺唔喺杯入便果枝筆比我。	
Q3. 擺離開盒最遠果個銀比我。	
Q4. 反轉個銀，俾枝鉛筆我，講比我聽你乜嘢名。	
Q5. 當我拍手嘅時候，你要擺起個波同埋掂吓自己個頭。	
Q6. 將個銀放係張紙中間。	

2. General Information

Age Group II

物件： 三個顏色盒及珠、鉛筆、波、六個銀、長方形盒

Q1. 床用嚟做乜嘢？	
Q2. 紅色珠擺係紅色盒，咁呢 D 珠應該擺係邊 D 盒呀？	
Q3. 天係上面、咁地係邊度呀？	
Q4. 厠所裏面嘅毛巾用嚟做乜嘢？	
Q5. 呢個乜嘢色？(四件物件)	
Q6. 耳仔用嚟做乜嘢？	

Age Group IV

物件：三個顏色盒、綠色車、一個銀

Q1. 厠所裏面嘅毛巾用嚟做乜嘢？	
Q2. 呢個乜嘢色？(四件物件)	
Q3. 耳仔用嚟做乜嘢？	
Q4.「消防員」做乜嘢架？	
Q5. 一隻橙切開一半，有幾多舊呀？	
Q6. 講比我聽你點樣洗手㗎？	

Age Group VI

物件：一個銀、鉛筆、波、紙

Q1. 「消防員」做乜嘢架？	
Q2. 一隻橙切開一半，有幾多舊呀？	
Q3. 講比我聽你點樣洗手㗎？	
Q4. 裙係乜嘢嚟架？	
Q5. 禮物係乜嘢嚟架？	
Q6. 放假係點樣架？	

3. Sentence Repetition

而家我地玩「小吱喳」遊戲。我會講 D 句子出嚟。我想你跟住我講。等我講完一句就到你勒。

Age Group II

Q1. 我有三個波。	
Q2. 隻狗有四隻腳。	
Q3. 哥哥好高。	
Q4. 呢架車係我嘅。	
Q5. 爸爸返咗屋企。	
Q6. 隻貓飲緊奶。	
Q7. BB 食曲奇餅。	
Q8. 我有一隻大狗。	
Q9. 個波係我嘅。	

Age Group IV

Q1. 個紅色波爆咗。	
Q2. 哥哥隻腳好污蹧。	
Q3. 佢地玩緊自己 D 玩具。	
Q4. 佢地想開生日會。	
Q5. 爸爸頂帽太大啦 ，我唔啱戴。	
Q6. 點解我唔可以出去呀？	
Q7. 我食咗早餐先返學。	
Q8. 哥哥著住條紅色褲。	
Q9. 個萍果擺係枱上便。	

Age Group VI

Q1. 個小朋友踩爛咗張凳。	
Q2. 佢做晒功課先可以睇電視。	
Q3. 媽咪煮緊鷄同埋豬扒俾我食。	
Q4. 我地食飯之前要洗手。	
Q5. 枝黑色筆跌咗落杯裏面。	
Q6. 爸爸做緊自己 D 嘢。	
Q7. 佢地踩緊佢地架單車。	
Q8. 佢頭先著好衫先出門口。	
Q9. 個小朋友係遊樂場度比佢哥哥推跌咗。	

4. Articulation

而家我地玩「跟我講」遊戲，
我會講 D 字出嚟，我想你跟我講。
我講「媽」， 咁你要講.....「媽」。
(Prompting : 你識唔識講呀？)
而家再玩過，但係今次我會遮住個嘴，好，我講.....

I. Onset

p	波		k	狗	
	包			鏡	
t	蛋		f	褲	
	豆			飯	
m	媽		ts	遮	
	門			豬	
j	月		s	手	
	雨			書	
l	路		p^h	婆	
	六			盤	
w	雲		kw	瓜	
	鑊			骨	
h	口		k^h	旗	
	蝦			橋	
n	你		t^h	頭	
	鳥			兔	
ŋ	我		ts^h	七	
	牛			乂	
			k^hw	規	
				跨	

II. Codas

p	葉	
	碟	
t	熱	
	八	
k	屋	
	黒	
m	三	
	心	
n	人	
	田	
ŋ	橙	
	水	
u	刀	
	貓	
i	雞	
	杯	

APPENDIX G
Written Consent of Parent/Guardian of the Subjects

敬啟者：

本人爲香港理工大學碩士研究生，亦爲在職職業治療師，現正進行一項有關兒童發展評估的研究，名爲 " Validation of Miller Assessment for Preschoolers (Cantonese Version)"，「米娜測驗」(Miller Assessment for Preschoolers) 是一個評估兒童發展的測驗，由一位美國職業治療師設計。是項研究的目的，是測試「米娜測驗」翻譯成中文後，對香港小朋友是否適用。希望邀請貴子女參加一項約四十五分鐘的測驗。在測驗中小朋友將會做一些大肌肉動作、認知概念及語言測試，測驗會作爲香港兒童發展能力的標準。如貴子女不願意參加，可隨時退出。所得資料，只作分析用途，並會絕對保密。

如閣下同意貴子女參加是項測驗，請填妥附上之同意書。

此致

貴家長

潘恩賜
理工大學碩士研究生
一九九七年七月四日

同意書

本人同意子女 _____ 〔子女姓名〕參加 " Validation of Miller Assessment for Preschoolers (Cantonese Version)" 此項測驗。並明白在研究過程中所得資料，只作分析用途，及絕對保密。本人之子女於測驗過程中，如不願意參加，可以隨時退出。

貴子女曾否在母嬰健康院進行三歲智能測驗 ：

無 ☐

有 ☐ 測驗結果： _____

_____
家長簽名

APPENDIX H

ALTERATIONS DURING BACKWARD TRANSLATION AND EXPERT PANEL REVIEWS (EQUIVALENCE)

Item	Original Version	Original Translation	Alterations	
			Backward Translation	Expert Panels
It 1	Big Building Game	砌高樓遊戲	砌高樓遊戲	砌高高遊戲
It 1	Make it as big as you can	盡量砌，砌得越高越好	可以砌到幾高就幾高	-
It 2	Now we are going to switch hands		我地而家換一換，用呢隻手嚟指	我地而家換過隻手嚟做
It 4	I don't make baby lines and I don't make giant lines		我唔會畫D好粗又或者好細條嘅線㗎	我唔會畫太長又或者太短嘅線
It 6	The Mr. Thumbuddy Game		大姆指叔叔遊戲	手指公叔叔遊戲
It 6	Now you try while we count to six		而家我由一數到六	而家我地連續做六次
It 8	Put one foot in front so they touch		擺一隻腳係前面掂住	擺一隻腳係另一隻腳前面
It 9	by accident		意外	唔小心
It 10	Can you roll up into a ball		成個波咁	捲埋好似個波咁
It 12	Move your tongue up		將條俐盡量向上郁	將條俐盡量向上伸

*It 1 - tower; It 2 - stereognosis; It 4 - motor accuracy; It 6 - hand-to-nose; It 8 - stepping; It 9 - walks line; It 10 - supine flexion; It 12 - tongue movement

Appendix I

Inter-item Correlation under three sub-scales of MAP: Foundations, Coordination, and Verbal

The inter-item correlation and respective significant level are shown. From Table 6.1 to 6.3 were the three age groups under foundations sub-scale. Table 6.4 to 6.6 were the three age groups under coordination sub-scale. Table 6.7 to 6.9 were the three age groups under verbal sub-scale.

It 1 - tower, It 2 -stereognosis, It 3 - finger localization, It 4 - motor accuracy, It 5 - vertical writing, It 6 - hand-to-nose, It 7 - Romberg, It 8 - stepping, It 9 - walks line, It 10 - supine flexion, It 11 - kneel-stand, It 12 - tongue movement, It 13 - rapid alternate movement, It 14 - follow direction, It 15 - articulation, It 16 - general information, It 17 - sentence repetition, It 18 - digit repetition.

Table 6.1

Inter-item correlation : Foundations Sub-scale (Age Group II)

Items	It 3	It 5	It 8.1	It 8.2	It 9.1	It 9.2
It 3	1.00					
It 5	.06	1.00				
It 8.1	.08	.22	1.00			
It 8.2	.16	.04	-.11	1.00		
It 9.1	.09	.06	-.14	-.12	1.00	
It 9.2	-.13	.02	.09	-.02	.08	1.00

Table 6.2

Inter-item correlation : Foundations Sub-scale (Age Group IV)

Items	It 3	It 5	It 8.1	It 8.2	It 9.1	It 9.2	It 10	It 13
It 3	1.00							
It 5	-.14	1.00						
It 8.1	.08	-.10	1.00					
It 8.2	-.01	.02	.05	1.00				
It 9.1	.10	.05	.14	.07	1.00			
It 9.2	-.15	-.15	-.06	-.18	-.15	1.00		
It 10	.32*	.08	-.21	-.04	-.11	.15	1.00	
It 13	-.03	-.06	-.25	.04	.29*	-.05	-.03	1.00

*$p < 0.05$

Table 6.3

Inter-item Correlation: Foundations Sub-scale (Age Group VI)

Items	It 2	It 3	It 5	It 7	It 8.1	It 8.2	It 9.1	It 9.2	It 10	It 13
It 2	1.00									
It 3	.09	1.00								
It 5	-.10	-.09	1.00							
It 7	-.04	-.03	-.04	1.00						
It 8.1	-.09	-.08	-.10	-.04	1.00					
It 8.2	.31*	.03	-.03	-.10	-.09	1.00				
It 9.1	.08	-.12	-.17	-.09	.07	-.06	1.00			
It 9.2	.41*	.12	-.09	.49*	-.08	.15	-.08	1.00		
It 10	-.08	-.07	-.09	-.03	.10	.03	-.19	-.07	1.00	
It 13	-.11	.03	.38*	-.05	.02	-.04	.12	-.10	-.10	1.00

* $p <$ 0.05

Table 6.4

Inter-item Correlation : Coordination (Age Group II)

Items	It 1	It 5	It 9.1	It 9.2	It 12	It 15
It 1	1.00					
It 5	.07	1.00				
It 9.1	-.25	.06	1.00			
It 9.2	-.10	.02	.08	1.00		
It 12	-.05	.11	.03	-.19	1.00	
It 15	.08	.13	-.03	-.09	.14	1.00

Table 6.5

Inter-item Correlation : Coordination (Age Group IV)

Items	It 1	It 4	It 5	It 9.1	It 9.2	It 12	It 13	It 15
It 1	1.00							
It 4	.01	1.00						
It 5	-.19	-.16	1.00					
It 9.1	-.10	-.11	.05	1.00				
It 9.2	.03	.07	-.11	-.15	1.00			
It 12	.20	.04	-.08	-.17	.13	1.00		
It 13	-.13	-.12	-.06	.29*	-.05	-.04	1.00	
It 15	.08	-.09	-.06	.25*	.03	-.09	.26*	1.00

* $p < 0.05$

Table 6.6

Inter-item Correlation : Coordination (Age Group VI)

Items	It 1	It 4	It 5	It 9.1	It 9.2	It 12	It 13	It 15
It 1	1.00							
It 4	-.11	1.00						
It 5	.05	-.04	1.00					
It 9.1	.05	-.09	-.17	1.00				
It 9.2	.32*	-.09	-.09	-.08	1.00			
It 12	.22*	-.11	-.04	.04	.03	1.00		
It 13	.06	-.05	.38*	.12	-.10	.14	1.00	
It 15	.27*	-.09	-.04	-.03	.22	.26*	-.05	1.00

* $p < 0.05$

Table 6.7

Inter-item Correlation : Verbal Sub-scale (Age Group II)

Items	It 14	It 16	It 17	It 18
It 14	1.00			
It 16	.57*	1.00		
It 17	.37*	.23*	1.00	
It 18	.41*	.24*	.98*	1.00

* $p < 0.05$

Age Group IV

Items	It 14	It 16	It 17	It 18
It 14	1.00			
It 16	.49*	1.00		
It 17	.12	-.01	1.00	
It 18	.19	.20	.32*	1.00

* $p < 0.05$

Age Group VI

Items	It 14	It 16	It 17	It 18
It 14	1.00			
It 16	.19	1.00		
It 17	.33*	.49*	1.00	
It 18	.23*	.00	.12	1.00

* $p < 0.05$

APPENDIX J
NORMATIVE TABLES OF CMAP

Table 7.1

Mean Scores of Hong Kong Children at Age Group II

Items	Mean Scores	S. D.	S. E. M.
tower	11.13	2.76	0.00
Stereognosis	3.93	0.25	0.00
finger localization	2.00	1.21	0.00
motor accuracy	5.77	3.40	0.00
vertical writing	5.93	3.06	0.60
hand-to-nose	5.74	1.05	-
Romberg	14.84	0.82	0.26
stepping (distance)	1.20	0.57	0
stepping (deviation)	1.95	1.01	0
walks line (speed)	11.33	4.61	0.33
walks line (error)	0.97	1.29	0.09
supine flexion	7.59	4.90	1.08
kneel-stand	1.97	0.26	-
tongue movement	2.48	0.91	0.10
rapid alternate movement	8.05	1.96	1.19
follow direction	2.31	0.85	-
articulation	30.46	2.94	0.67
general information	2.15	0.68	-
sentence repetition	2.57	0.83	-
digit repetition	2.57	1.22	0

Table 7.2

Mean Scores of Hong Kong Children at Age Group IV

Items	Mean Scores	S. D.	S. E. M.
tower	11.68	2.75	0.00
Stereognosis	3.93	0.25	0.00
finger localization	2.68	0.91	0.00
motor accuracy	8.83	3.40	0.00
vertical writing	6.25	3.05	0.59
hand-to-nose	5.93	0.52	-
Romberg	15.00	0.00	0
stepping (distance)	1.20	0.61	0
stepping (deviation)	1.67	0.95	0
walks line (speed)	10.92	6.02	0.43
walks line (error)	0.47	1.69	0.12
supine flexion	11.68	4.64	1.02
kneel-stand	2.00	0.00	-
tongue movement	3.00	0.97	0.10
rapid alternate movement	6.83	1.61	0.98
follow direction	2.13	0.83	-
articulation	31.50	2.90	0.66
general information	2.68	0.57	-
sentence repetition	1.9	0.90	-
digit repetition	3.67	0.88	0

Table 7.3

Mean Scores of Hong Kong Children at Age Group VI

Items	Mean Scores	S. D.	S. E. M.
tower	12.84	2.29	0.00
Stereognosis	3.67	0.62	0.00
finger localization	3.41	0.68	0.00
motor accuracy	10.52	2.99	0.00
vertical writing	4.31	2.14	0.42
hand-to-nose	5.98	0.12	-
Romberg	13.50	3.20	0
stepping (distance)	1.17	0.58	0
stepping (deviation)	1.53	0.71	0
walks line (speed)	28.73	8.93	0.64
walks line (error)	2.69	3.17	0.23
supine flexion	13.86	3.08	0.68
kneel-stand	2.00	0.00	-
tongue movement	3.38	0.79	0.08
rapid alternate movement	6.16	1.28	0.78
follow direction	2.27	0.84	-
articulation	32.64	1.38	0.31
general information	2.03	0.94	-
sentence repetition	2.05	0.86	-
digit repetition	4.27	0.70	0

Table 7.4

Item Difficulty and Discrimination of Hong Kong Children: Foundations Sub-scale

Items	Item Difficulty			Item Discrimination		
	II	IV	VI	II	IV	VI
stereognosis	0.98	0.98	0.92	0.15	0.16	0.43**
finger localization	0.60	0.54	0.85	0.38**	0.59**	0.10
vertical writing	0.72	0.80	0.72	0.41**	0.48**	0.40**
hand-to-nose	0.96	0.99	0.99	0.15	0.23	-
Romberg	0.99	1.00	0.90	0.18	-	0.37**
stepping (distance)	0.95	0.95	0.76	0.03	0.18	0.22
stepping (deviation)	0.81	0.87	0.87	0.34**	0.37	0.35**
walks line (speed)	0.70	0.58	0.49	0.14	0.33*	0.04
walks line (error)	0.84	0.96	0.76	0.37**	0.17	0.55**
supine flexion	0.54	0.73	0.92	0.69**	0.54**	0.34*
kneel-stand	0.99	0.67	1.00	0.19	-	-
rapid alternate movement	0.43	0.28	0.36	0.45**	0.36**	0.42**

**$p<0.01$, *$p<0.05$

Table 7.5

Item Difficulty and Discrimination Indices of CMAP : Coordination Sub-scale

Items	Item Difficulty			Item Discrimination		
	II	IV	VI	II	IV	VI
tower	0.70	0.73	0.80	-0.04	0.24	0.36**
motor accuracy	0.45	0.55	0.55	-0.13	0.50**	0.52**
vertical writing	0.72	0.80	0.72	0.41**	0.53**	0.51**
walks line (speed)	0.70	0.58	0.49	0.14	0.36**	0.44**
walks line (error)	0.84	0.96	0.76	0.37**	0.18	0.12
tongue movement	0.62	0.80	0.84	0.08	0.45**	0.51**
rapid alternate movement	0.43	0.28	0.36	0.45**	0.31**	0.34**
articulation	0.90	0.93	0.96	0.10	0.35**	0.05

**$p<0.01$, *$p<0.05$

Table 7.6

Item Difficulty and Discrimination of Hong Kong Children : Verbal Sub-scale

Items	Item Difficulty			Item Discrimination		
	II	IV	VI	II	IV	VI
follow direction	0.83	0.78	0.82	0.68**	0.67**	0.45**
general information	0.79	0.67	0.76	0.61**	0.46**	0.64**
sentence repetition	0.90	0.72	0.77	0.52**	0.66**	0.71**
digit repetition	0.52	0.73	0.85	0.75**	0.58**	0.70**

**$p<0.01$

APPENDIX K

INSTRUCTION SHEET FOR FIELD TEST

1. 砌高高遊戲
   我地而家玩「砌高高」遊戲。
   你可唔可以砌座高樓好似我呢座咁高呀？
   可以砌到幾高就幾高。
   (Prompt : 你可唔可以砌多一「舊」上去呀？)

2. 摸吓估吓遊戲
   而家我地玩「摸吓估吓」遊戲
   呢 D 玩具係你嘅。(擺開兒童盒內物件)
   呢 D 玩具就係我嘅。(考官物件連盒放到「牆」後)
   我同你 D 玩具係一樣嘅。
   我地當呢幅係牆，
   我將一件玩具放係呢隻手度。(指左手)
   你就用呢隻手(右手)指番同一件玩具出嚟。
   我會擺幅牆係度，等你睇唔倒。(將「牆」放在兒童左手手腕上)
   準備好未？呢個係邊件玩具呢？
   (Prompt : 呢隻手要放係牆後面架，用呢隻手嚟指。)
   我地而家換過隻手嚟做。
   你用呢隻手嚟摸件玩具，用呢隻手嚟指。

3. 手指指遊戲
   我地玩一個叫「手指指」嘅遊戲，將你隻手咁擺。
   我而家掂你隻手。(中指)
   我掂你邊度呀？你用呢隻手指俾我睇。
   我而家擺幅牆係度，等你睇唔倒。
   指俾我睇我掂你邊度。
   準備好未？
   我掂你邊度呀？

4. 畫雀仔籠 (限時 20 秒)
我地而家玩一個「畫雀籠」嘅遊戲。
我將會係呢度畫個雀籠。睇住我。(示範)
睇吓，我好小心畫架。
我 D 線長短啱啱好架。
佢唔會穿過底，又唔會穿過頂。亦都唔會好彎曲。
(一面畫、一面搖頭)
你係呢度練習畫吓 D 線先，
跟住你就可以畫個大雀籠。係呢度試吓。
(或: 我唔會畫太長又或者太短嘅線。)
畫得越多線越好。直到我叫你停爲止。
準備好未？開始！
要好小心喎，試吓畫快 D 。
好，停嘞。

5. 兔仔跳回家
我地而家玩「兔仔返屋企遊戲」。
我地要帶兔仔跳返屋企，睇住我啦。(示範)
跳，跳，跳
而家兔仔跳到去間屋度嘞。
而家到你做，由兔仔果度開始。
記住隻手要遞高喎。
帶隻兔仔跳返屋企。跳，跳，跳。
我地再試多次。記住要遞高隻手喎。(可以幫兒童放好手的位置)
我地試吓可唔可以唔睇住嚟做？
我舉起呢幅牆，等你睇唔倒。
準備好未？
跳，跳，跳。繼續啦。

6. 手指叔叔遊戲
我地一齊玩一個「手指叔叔」遊戲。
我將隻手擺個「停止」嘅手勢。
首先手指叔叔掂我隻手，跟住掂我個鼻。
然後佢再掂我隻手，再掂我個鼻。
好，而家輪到你嘞，你想用邊隻手呀？
(將公仔放在兒童的食指上)
而家要手指叔叔掂我隻手，
好，掂你個鼻。
跟住掂我隻手。做得好好喎！
而家我地連續做 6 次，
你要用手指叔叔掂我隻手，跟住掂你個鼻。
而家試吓合埋對眼嚟做。
睇吓手指叔叔可唔可以掂倒我隻手跟住掂你個鼻，一路都要合埋對眼㗎，我一路數，睇你可唔可以做到六吓？

7. 扮石像遊戲

Age Group I-IV:
而家我地玩一個「扮石像」遊戲。
將對腳拍埋好似我咁。你試吓。
而家睇你可唔可以合埋對眼企係度唔郁，
直至我數到第十五吓爲止。
企定係度好似個石像咁。

Age Group V-VI ：
我地而家玩一個「扮石像」遊戲。
我將對腳放係哩 D 腳印上面，好似咁。
而家你試吓，擺一隻腳係另一隻腳前面，
要腳掂住腳。
睇吓你可唔可以合埋對眼企係度，
直至我數到十五吓爲止。
好，企定係度好似個石像咁唔郁，合埋對眼。

8. 操兵遊戲 (限時 20 秒)
而家我地玩一個操兵遊戲。
你企係呢 D 腳板上面，
跟住好似我咁踏步，盡量留係 D 腳板上面喎。
而家試吓合埋眼操，操到我數到二十爲止。
對腳盡量踏係腳板上面。
(Prompting：我地要操去一個好遠嘅地方，
一二，一二，一二，咁操上山。
繼續一二，一二，一二，
我地就快操到屋企啦，一二，一二，一二，
經過間士多，一二，一二，一二.....)
好，停嘞。

9. 行直線遊戲
我地而家玩「行直線」遊戲。

Age Group I-IV:
你可唔可以好似我咁做呀？睇住我啦。
如果我唔小心踏出條線，
我要將隻腳放番係條線上面。
你試吓行，對腳要踏住條線，唔好踏出喎。
由呢度開始，你盡快行到去尾，
但係唔好踏出條線喎。
如果你唔小心踏出條線，
就要將隻腳放番係條線度。
而家由頭再試一次，
準備好未？開始。

Age Group V-VI:
我地一齊試吓係條線上面咁樣行，睇住我。
我放低呢隻腳，腳踭要掂住腳尖。
你係我後面跟住我行。
而家係呢度開始，
盡量踏住條線行，如果唔小心踏出條線，
將隻腳踏返上條線度。
睇吓你幾快可以行到尾，而又冇踏出條線。

10. 扮波波遊戲
我地而家玩「扮波波」遊戲。
你可唔可以捲埋好似個波咁呢？睇住我啦。(示範)
背脊瞓係度，
屈埋對手同對腳好似咁，
睇吓你個頭可以掂住小丑個頭幾耐，
我會由一數到十五 ， 睇吓你可以保持幾耐，
(Prompting : 掂住，掂住。)

11. 起立遊戲
而家我地玩一個「起立」遊戲。
你可唔可以跟我一模一樣咁做？
睇住，而家企起身。

12. 俐仔郁郁遊戲
我地而家玩「俐仔郁郁」遊戲。
你條俐可唔可以好似我咁郁呀？
將條俐盡量向上伸。

13. 踏踏腳遊戲
而家我地玩個「踏踏腳」遊戲。
你對腳可唔可以好似我咁踏呀？
睇吓你可以踏得幾快先。

14. 我講你做

我地而家玩一個「我講你做」嘅遊戲。
你聽清楚我叫你做乜，你就照做。
拍手掌。(Prompting:你識唔識拍手掌呀？/示範拍手；好，拍得好。)
我可能會一次過叫你做幾樣嘢。
等我講完之後，你先可以做喎。
* If the child begins before you are finished, say:
「我未講完喎，等我講完你先至去做啦。」(每個動作限時 10 秒)

Age Group II

物件： 一個盒、三枝筆 (不同長短) 及 波

1. 係盒度拎兩枝筆出嚟。	
2. 係枱度搵枝筆比我。	
3. 擺果枝筆係盒側邊。	

Age Group IV

物件： 五個硬幣、一個盒、三枝筆(不同長短)、一架車(頭向小朋友)

1. 擺晒 D 筆同埋銀係盒裏面。	
2. 擘開個口， Lam 吓你個口唇，然後笑一笑。	
3. 將個銀放係車仔後面。	

Age Group VI

物件： 五個硬幣、一個盒、三枝筆(不同長短)

1. 擺短果枝筆喺杯側邊。	
2. 攞唔喺盒入便果枝筆比我。	
3. 攞離開盒最遠果個銀比我。	

15. 跟我講

而家我地玩「跟我講」遊戲，
我會講 D 字出嚟，我想你跟我講。
我講「媽」，咁你要講.....「媽」。
(Prompting : 你識唔識講呀？)
而家再玩過，但係今次我會遮住個嘴，好，我講.....

p	波		ts	遮	
	包			豬	
t	蛋		s	手	
	豆			書	
m	媽		p^h	婆	
	門			盤	
j	月		kw	瓜	
	雨			骨	
l	路		k^h	旗	
	六			橋	
w	雲		t^h	頭	
	鑊			兔	
h	口		ts^h	七	
	蝦			乂	
k	狗		k^hw	規	
	鏡			跨	
f	褲				
	飯				

16. 一般知識

Age Group II

1. 床用嚟做乜嘢？	
2. 厠所裏面嘅毛巾用嚟做乜嘢?	
3. 天係上面、咁地係邊度呀？	

Age Group IV

物件：四個顏色盒

1. 呢個乜嘢色？(四件物件)	
2. 耳仔用嚟做乜嘢？	
3. 消防員做乜嘢架？	

Age Group VI

1. 講比我聽你點樣洗手㗎？	
2. 裙係乜嘢嚟架？	
3. 禮物係乜嘢嚟架？	

17. 小吱喳遊戲

而家我地玩「小吱喳」遊戲。
我會講 D 句子出呢。
我想你跟住我講。
等我講完一句就到你喎。

Age Group II

1. 我有一隻大狗。	
2. 個波係我嘅。	
3. 呢架車係我嘅。	

Age Group IV

1. 佢地玩緊自己 D 玩具。	
2. 佢地想開生日會。	
3. 點解我唔可以出去呀？	

Age Group VI

1. 個小朋友踩爛咗張凳。	
2. 佢做晒功課先可以睇電視。	
3. 枝黑色筆跌咗落杯裏面。	

18. 數字遊戲

我地而家玩數字遊戲。

我會讀 D 數字出嚟。

我讀乜你就跟我一模一樣讀乜。

等我讀完就輪到你勒。

準備好未。

4, 2	
5, 1, 6	
8, 4, 1, 7	
9, 6, 2, 5, 1	
8, 3, 4, 1, 6, 9	