

**CITY UNIVERSITY OF HONG KONG**

Development and Validation of a Quick Screening Test for Local Hong  
Kong Cantonese-speaking Intellectually Disabled Adults: a 10-item  
Checklist for use by non-professionals.

by

Edward Lau Yun Leung

SS5790 Research Paper

Submitted to Department of Applied Social Studies

in Partial Fulfillment of the Requirements for the

Post Graduate Diploma in Psychology

May 2007

# *Acknowledgements*

I would like to express my sincere gratitude to my dissertation supervisor, Dr. John Tse, for giving me definitive direction and much-needed support. Dr. Tse is unbelievably patient, sharp-minded, and knowledgeable in the field of Applied Psychology, especially in the Intellectual Disability domain. I feel most fortunate to have been under his supervision and have been the beneficiary of his unequivocal opinions and guidance, none more so than his advice for this topic choice.

I also wish to thank Father Bonzi, Mrs. Wong, Emily Lam, Spencer Suen, Joseph Au, Eric Siu, Ng kin-wah, Kitty Ho of Fu Hong Society, Mrs. Cheuk and Macy Law of the Joint Council of parents of mentally handicapped people, Chan ping-leung and Patrick S.H. Wong of St. James Settlement, Bok kam-lung of Wi-Chi Service, Emily Fung, tai-hong and all my Chosen Power friends, Tong siu-hong of Caritas, and the various workshop staff for their support, smiles, and tea and coffee. They provided our research team with the essential research target subject sources and most hospitable facilities which have helped the research process enormously.

Unreserved gratitude is also extended to all the participants, siu-wan, Mr. Walter Yu , Ms. Mui yee-wai and other Focus Group members who have spent time and patience catering for the research needs to provide me with the valuable research information that have made this project possible.

I dedicate this to Winly.

## *Abstract*

Previous research to develop quick screening tests for intellectually disabled (ID) people has shown over-reliance on physical characteristics identification or complex apparatus set-up when designed for professional administration. This study examined another mode of operation, making use of mainly adapted conservation task-items of freely-available local coins and bank-notes set-up, with the intention to place such tools in the hands of a layman to check and assist his fellow ID adult citizen, in case of emergency missing/ lost scenarios and during work or leisure interactions.

The 68 adult ID participants, 26 females and 42 males, sought out by purpose sampling, in a rough 4:3:1 moderate/mild/borderline ratio, were administered the 13 items through friendly standardized procedures at their work or activity sites with face validity obtained through 2 Previous Focus Groups, and ecological validity checked through prior rapport-building semi-structured interviews. Total score means were statistically *t*-tested with those of 2 comparison groups of 50 MA-matched non-disabled children and 50 CA-matched non-disabled adults.

Item analysis of item-difficulty (*p*) and item-discrimination (*D*) values helped finalize a 10-item scale with desirable high *p*- (.06 to .37 except one as .53) and *D*- (.54 to .78) values, and Cronbach Alpha of .91 and Split-half coefficients at .91 to .97. Statistical analyses showed significant total group score means *t*-test differences but no gender bias.

Results were discussed on the practical functionality of whole scale and its partial usage in terms of simplicity, speed, and differentiating and motivating characters. Methodological clumsiness and possible scale misuse were spelled out as limitations. Future research was suggested for betterment.

## TABLE OF CONTENTS

<u>Content</u>	<u>Page no.</u>
Acknowledgement	ii
Abstract.....	iii
Table of Contents.....	v
List of Tables.....	viii
<b>Chapter 1: Introduction and Literature Review.....</b>	<b>1</b>
1.1 Introduction.....	1
1.2 Literature Review.....	5
1.2.1 ID: Contemporary Views according to AAMR, ICF, and DSM-IV-TR.	5
1.2.1.1. Adaptive Behaviour Criterion & Representative Scales.....	6
1.2.1.2. ID Intellectual Functioning Criterion & Representative indicators (IQ and MA).....	7
1.2.2. ID: According to Inhelder’s Operational Diagnosis of Reasoning.....	8
1.2.2.1 IQ & Inhelderian Operational Indicators Compared.....	12
1.2.2.2. Links between MA, IQ, & Inhelderian Operativity Scores.....	13
1.2.2.3 Additional Research for Cognitive Indicators for children aged 9 -11.....	14
1.2.3. Quick Screening Tests Reviewed.....	14
1.2.3.1. Quick IQ-Substitute Tests.....	14
1.2.3.2. Quick Cognitive Functioning Tests: MMSE.....	15
1.2.3.3. Quick Screening Tests for non-professional use.....	16

1.2.3.4.	Difficulties in Developing Quick Screening Tools.....	17
1.2.4.	Selection of Focus Group Discussion Domains.....	18
1.2.4.1.	Selection of Intellectual Functioning Domains as “Cognitive Functioning” Domains.....	18
1.2.4.2.	Selection of Adaptive Behavioral Domains as “Daily Common Sense” Indicators.....	19
1.2.5.	Criteria for Task Inclusion.....	20
<b>Chapter 2:</b>	<b>Methodology</b> .....	<b>22</b>
2.1	Research Design.....	22
2.2	Participants.....	22
2.3.	Apparatus & Materials.....	23
2.3.1.	Apparatus & Materials for Focus Group Discussion.....	23
2.3.2.	Apparatus & Materials for Interviews with ID adult participants....	24
2.3.3.	Apparatus & Materials for Scale Administration.....	24
2.4	Procedures.....	26
2.4.1.	Procedures for Focus Group Meetings.....	26
2.4.2.	Procedures for Interviews with Target Group Participants.....	26
2.4.3.	Procedures for Scale Administration.....	27
2.5.	Data Analysis.....	28
<b>Chapter 3:</b>	<b>Results and Findings</b> .....	<b>29</b>
3.1	Item Analysis.....	29
3.1.1	Split-half Reliability Checks.....	32
3.1.2	Gender Differences.....	34
3.2.	Across-group Comparison of Scale Item Scores.....	34
<b>Chapter 4:</b>	<b>Discussion</b> .....	<b>35</b>
4.1	Summary and Discussion of Findings.....	35

4.1.1. Scale Reliability and Item Statistics Findings.....	35
4.1.2. Scale Performance Across Target and Comparison Groups.....	36
4.1.3. Gender Bias.....	36
4.1.4. Practical Considerations for Individual Item Functionality.....	36
4.2. Implications.....	39
4.2.1. Theoretical Implications.....	39
4.2.2. Implications for Our Research Objective.....	40
4.2.3. Implications for Possible Misuse.....	40
4.3. Methodological Limitations.....	41
4.4. Cross-cultural Concerns.....	42
4.5. Suggestions for Future Research.....	44
4.6. Conclusion.....	44

References

Appendix A

Appendix B

Appendix C

Appendix D

Appendix E

Appendix F

## **LIST OF TABLES**

<b><u>Tables</u></b>		<b><u>Page no.</u></b>
Table 1	Item Statistics of Item Difficulty for ID Adult Group Total Score Means sorted from Hardest (lowest <i>p</i> -values) to Easiest (highest <i>p</i> -values) Items..	30
Table 2	Item-Total Statistics in order of descending D-values to explore and discard low D-value Items.....	31
Table 3	Finalized 10-item Scale: Item-Total Statistics in order of Descending D-values.....	32
Table 4	Summary of Split-half coefficients and Inter-form Correlations of Exploratory Split-half Combinations.....	33



# *Chapter 1: Introduction and Literature Review*

## *1.1 Introduction*

How do people with intellectual disabilities reason, through their daily challenges in a hasty-rush community like Hong Kong? News of local intellectually disabled (ID; abbreviation used throughout this report) adults gone lost like 45-year-old Kwok (Ming Pao, 2007, February 26) and 59-year-old Lai (Ming Pao, 2007, May 7) seems infrequent and yet sometimes carries hurtful consequences for their worrisome family members, if not identified by physical appearances by police and the public early enough. Both afore-mentioned incidences occurred when the disabled citizens' commute routines were broken, one due to fire-induced road blockage, the other sudden loss of visual contact with care-assistant just after alighting transit vehicle. While lost, Kwok went around on foot for 30 hours while Lai on bus-rides for 48, both spent lonely sleeping hours in parks, and only recovered by two photograph- alerted police officers and one public citizen respectively.

Seventeen-year-old ID youngster Lai cut her neck to commit a bloody suicide (Ming Pao, 2007, April, 21) after taking her parents' threatening words to throw her into government-run detention homes (for repeated wrist-cutting) literally that it took the rare combined effort of seven worn-out police negotiating specialists receiving

constant phone-linked psychologist advices to try reasoning with her and resorting to diversion tactics in the end to disarm and 'rescue' her.

Other ID adults, it seems, when faced with challenges but supported accordingly, are motivationally and physically fit enough to integrate with local non-disabled citizens in joining demanding social activities like the ten ID swimmers participating in the second Hong Kong 3,000 meters swim safety campaign (Tsang, 2007, May 2) or, as 21-year-old Lam and 29-year-old Wu, both Downs people, earning praise for making successful breakthroughs in landing government jobs (Hu, 2007, March 21).

Clearly, local ID people vary in intellects, personalities, and affects just like our typical population, only sharing amongst each and every one of them a common statistically marked two standard deviations or more downward shift of Intelligence Quotient score established by clinical psychologists for educational and vocational placement purposes.

The present study aims to produce scale task-items that have no reliance on physical appearance stereotyping, and of a non-interviewing questioning format through a systematic literature review for psychological ideas suitable and testable by psychometric principles for reliability and validity. But, why put such knowledge and its application in the minds and hands of a common man?

Very simply, the public is the vast majority in society and the man or woman in the street has social obligations not fulfilled unless properly informed. So, how informed is the average Hong Kong person for the welfare of his one-in-fifty (Hong Kong Government, 1999) ID neighboring citizens, amid government and academic urges for positive social integration? In a most admirable response to the need to pursue ecological knowledge to deepen the general public's understanding of what ID entails and first-hand social interaction behaviors of the intellectually disabled, Tong, Chan and Yau (2002) made painstaking efforts to produce a 9-month endeavor of a literal interaction mix of 20 ID trainees with non-disabled students from 3 primary and 1 secondary local schools and senior citizens residing in community-homes through 18 nuclear activities. Sharing their promotional concerns for ID people's social integration as an interpersonal phenomenon, our project was initiated directly in response to a recent government directive (Social Welfare Department, 2005) issued to raise awareness of ID citizens' wander-and-lost issue for immigration officers after the sad Yu man-horn incidence (Equal Opportunities Commission, 2001). Integration is comprised of three components, knowledge – mutual understanding, attitude – mutual acceptance, and behavior – interaction (Tong, Chan and Yau, 2002). We aim to explore the cognitive and adaptive behaviors of adult citizens disabled by with intellectual limitations to provide the common man, including disciplinary officers,

teaching instructors, work supervisors, arts and crafts or physical activity leaders, or rehabilitation personnel a supportive tool for his ID students, workmates, clients or friends who need help and support necessitating quick understanding of intellectual limitations primarily for immediate physical safety concerns and secondarily any communicative assistance in their study, work, play or sickness. There is a definitive attempt to put aside hitherto reliance upon physical appearance as sole ID status identification tool as help-conscious citizens (and police officers and as non-professionals alike) practiced for recovering the afore-mentioned ID adults.

What we are looking for are simple-to-set-up tasks requiring absolutely freely-available money items and writing utensils that could elicit markedly different behavioral responses from ID adults and non-disabled adults.

We set out with two assumptions. One is that a local screening test should have three components as “cognitive ability”, “daily common sense”, and “general knowledge” as achieving content-related evidence of validity. The other being that, to tap distinct intellectual or adaptive behavior indicators, reference has to be made to empirically established ID criteria (reviewed as follows) that allows systematic selection of intellectual and adaptive behavior domains in psychometrically sound Scales that could be integrated with our local contextual component elements through expert consensus to achieve ecological validity for the component items (e.g. the

component “*general knowledge*” as wholly grounded in focus-group opinions).

## ***1.2 Literature Review***

ID is a functional disability, not a medical or mental disorder (Luckasson et al, 2002), a disability of intellectual capacity that a person carries through life (Hogg & Langa, 2005).

ID is increasingly favored to replace outdated terms like mental handicap or mental retardation (Hogg & Langa, 2005; Kwok, 1997). ID is used here with its levels of cognitive impairment same as mental handicap grouping traditions as mild, moderate, severe, and profound.

### ***1.2.1 ID: Contemporary Views according to AAMR, ICF, and DSM-IV-TR***

The most updated and specific disability–orientated classification is that of the 2002 American Association on Mental Retardation (AAMR) system that defines ID socially. Its central focus is on social adaptation to hit home that ID adults with low intelligence quotient (IQ; abbreviation used throughout this report) levels could still function adequately by being socially skillful or aware enough (Drew & Hardman, 2007). Hence, emphasis is on significant limitations both in intellectual functioning and adaptive behavior, the latter further expressed as conceptual, social, and practical

adaptive skills, such disability further defined as present early in life and before age 18 (Luckasson et al, 2002).

Furthermore, the AAMR definition sees ID as multidimensional and is affected positively by individualized supports. It focuses on typical rather than maximal behavior, and includes subjective aspects of functioning; complimenting the professional, objective view of functioning and disability of the International Classification of Functioning, Disabilities, and Health (ICF) (WHO, 2001). Its three dimensions of adaptive behavior have strong empirical support through factor-analytic studies and conceptual work (Greenspan, 1999, Greenspan & Driscoll, 1997; Gresham & Elliott, 1987; Thompson et al., 1999; in Luckasson et al, 2002). It has very clear statistical diagnostic criteria for both IQ and adaptive behavior as follows.

#### ***1.2.1.1. ID Adaptive Behavior Criterion and Representative Scales***

Significant limitations in adaptive behaviors are operationally defined as performance that is at least two standard deviations below the mean of one of the three behavior domains in the presence or absence of an overall score on a standardized measure of all three domains.

The Vineland Adaptive Behavior Scale-Survey Form (Sparrow, Balla, & Cicchetti, 1984) enjoys good reliability and validity data support, and was normed upon ID and general populations and used in a fittingly conversation format (but “indirectly” administered to care-givers of ID targets).

A locally developed Home-based Training Package (HBTP) for ID school-leavers (Crawford & Tse, 1988) was included for its local relevance.

#### ***1.2.1.2. ID Intellectual Functioning Criterion and Representative Indicators***

##### ***-----IQ & MA (Mental Age)***

The intellectual functioning criterion for diagnosis of ID is approximately two standard deviations below the mean, (considering the SEM for the specific assessment instruments used and the instruments’ strengths and limitations) (Luckasson et al, 2002, 58) which are comparable to the Diagnostic and Statistical Manual’s (DSM-IV-TR) (American Psychiatric Association, 2000) recommendations regarding IQ cutoff scores.

IQ is currently relied upon heavily as a global, general factor as the universal measure of intellectual functioning (Luckasson et al, 2002). Despite a ease abuse for quick stereotyping when taken as a utility concept, and that, when commonly communicated as a convenient summary as a single score, it reveals little information

about specific skills, and even as separate subset scores, tell only quantitative representation of individual differences, its strength lies in its strong factorial support and relative minimal age-dependency when used contemporarily as **deviation IQ**, which is derived statistically through standard deviation computation and referenced to a standard score. Even with its reductionist and labeling risks, and unsatisfactory representation of human intelligence (Luckasson et al, 2002), deviation IQ remains the dominant ID impairment level discrimination index for educational and vocational placement. (The original **ratio IQ** calculation approach by dividing a person's mental age by his chronological age faded out about the mid-1940s because **ratio IQ** was unstable across age-levels and trouble for people approaching adulthood and) (Drew & Hardman, 2007).

MA is a construct measuring children's developing intellect, and is the operational definition of developmental levels (Weisz & Yeates, 1981). It is norm referenced to normal children's chronological-age-related behaviors and so quite easy to grasp by parents and laypeople for comparing ID persons, though again, being a summed figure it tells little about specific skills or deficits. As children are still developing, mental age is unstable and valid only for adults.

### ***1.2.2. ID: according to Inhelder's Operational Diagnosis of Reasoning***



Inhelder, in her pioneering work “The diagnosis of reasoning in the mentally retarded” (1943, 1968, cited in Paour, 2001), proposed a psychological definition of ID, that was based on her homogeneity of reasoning research. It was a cross-sectional study of 159 ID and disabled children and adults who were aged 7,5 to 52 (mean=13.2, S.D.=5,5); with IQ 42 to 105 ( mean = 73.7, S.D.= 11.5). Inhelder considered her reasoning levels a central measure of intellectual potential (Paour, 2001).

Her “functional” assessment used open-dialogue clinical interviewing methods and tasks. The tasks were of two types: substance, volume, weight conservation and weight transitivity to observe logical (concrete operation) reasoning; and sugar dissolution to check empirical and factual knowledge (Paour, 2001).

Her results showed a distribution of participants among 7 distinguishable levels of “completed” stages (88%) and “intermediate” stages (12%).

Paour (2001, p. 25) recorded Inhelder’s proposed ‘new’ criteria of ID and its degrees that reflect the stages of operational development (1943/1968, p. 293):

*...mental deficiency begins when the subject will never be able to make up his retardation of operatory construction [italics in original]. 1. the idiot never outgrows the sensory-motor compositions (previous to language). 2. the imbecile is capable of intuitive thought (egocentrism, irreversibility, but no operation). 3. the retardate is capable of operatory construction which he is incapable of completing,*

*i.e., “concrete operations” as opposed to formal operations and thus in time of reaching the full development of a normal child.*

“Idiot” denotes profound ID (IQ < 20), “imbecile” denotes severe and moderate ID ( IQ 20-49), while “retardate” means mild ID plus borderline cases (IQ 50-80) (Clarke & Clarke, 1974). Such obsolete ID nomenclature still let through Inhelder’s definitively detailed adaptation of Piaget’s staged ideas to lend us the rationales for positive discriminative indices that could elicit quick and obvious task responses from our target ID adults that, being comparable to non-disabled children, would very likely to be different from non-disabled adults who should have achieved formal (logical) operation or at least be fully capable of concrete operations (Weiten, 2004). The Piagetian developmental sequence as common textbook knowledge (Weiten, 2004, p.441-p.444) states that normal children aged 0 to 2 as in the sensory-motor stage, those aged 2 to 7 as in the pre-operational stage, 7 to 11 as in the concrete-operational stage, and beyond that, formal logical thought.

Worth mentioning for its limitations, Inhelder’s inference that her data could be referred to as an explicit and coherent developmental hierarchy was undermined by her non-standardized procedures and incomplete analysis (Weisz & Zigler, 1979), and, as the ultimate inherent weakness of all Piagetian operational “mental level indices”, her reasoning level data, needed MA correction for developing children that forbid

one unified scalogram construction for all age groups to challenge the IQ notion as a comprehensive intelligence measurement index (Paour, 2001).

As such, reasoning diagnosis is faulted by its age-dependency as a comprehensive intellect measuring tool for developing ID and non-ID children alike. This led Paour (2001) to judge Inhelder's work as "the practical failure of a theoretical success" (p. 26).

Inhelder's criteria had never been used for diagnostic purposes but Paour's (2001) emphasis on reasoning levels as a developmental guide to have found usefulness as a target-designating cognitive indicator (employed for planning psycho-pedagogical interventions) builds confidence for our screening tool. On closer examination, Inhelder's data did succeed in demonstrating a very important fact that reinforced earlier research findings that cognitive development of ID individuals do follow the same sequences, as later confirmed by Weisz & Zigler (1979)'s meta-analysis (that neither homogeneity nor similar sequence had been challenged). Paour's (2001, 19) elaboration "that mentally retarded subjects use reasoning entirely homologous with that of younger, non-retarded children" state a blatant relevancy to the present project. Hence, thus literature provides genuine theoretical and practical support for our adaptation of these Piagetian number, relation, and else physical quantity domains (e.g. length and weight) to test our adult ID targets which could

materialize into task-items just like Cinderella's feet-for-shoes adventure if they find their ways to retro-fit our quick screening task apparatus limitations as we shall see now.

Inhelder's work attracted many replication studies using various operational tests exploring numerous conceptual domains (number, relations, and classes; spatial relationships, space, time, speed; chance; logic; moral development) (Paour, 2001). Amongst which, we found conservation tasks of number, length and seriation tasks most optimally depicted by handy coins and banknotes and freely-available A4 paper that could be rolled up as paper strips. Katz & Ziv (1992) alerted us for feasible use of verbal instructions (stand on one foot or bend your knees and bow down) to test testees' own body weight as weight conservation test of their occupational therapy assessment of psychiatric patients.

Still, in order to establish validity references of possible reasoning (conservation or seriation) tasks, we need to compare diagnostic capabilities between Inhelderian levels with other established cognitive indicators like IQ and MA.

#### ***1.2.2.1 IQ and Inhelderian Operational Indicators Compared***

Zarro (1973, cited in Paour, 2001) compared IQ and Inhelder's reasoning level data and showed, after correcting reasoning data of younger subjects with their MA (as the levels of reasoning in younger subjects may still progress), that reasoning

levels led to harsher diagnosis than IQ. (He was, of course, vividly demonstrating the limitation of Inhelder's approach, which, like MA, operational diagnosis is only valid for adults).

#### ***1.2.2.2 Links between MA, IQ, and Inhelder's Operationality Scores***

Pasnak, Wilson-Quayle, & Whitten (1998) used 3 Piagetian conservation tasks (conservation of discrete quantities, seriation, and categorization) to confirm significant correlations between Inhelder's 1 to 7 grading (of conservation levels) with IQ (.66) and MA (.77). Paour (2001) used similar correlation analysis on Inhelder's data to find that operationality scores increases as a function of MA, thus confirming Inhelder's reasoning diagnosis as a developmental indicator. This would become the empirical basis of our Piagetian tasks for MA and CA (Chronological age) matching. Adjusting established age-level recommendations from ICF (Who, 2001), we hence set MA-matching age-levels for our moderate, mild, borderline ID adult as 6 to 8.9, 9-11.0, and, 12 and beyond.

In the same exercise, Paour showed 31 participants (out of 100) in the IQ range of 65-84 (i.e. the least mildly ID and the even better-graded borderline ID individuals) achieving MA levels of 10 and 11. As a precaution to safeguard our screening tool's ability to not let through thus best-performing ID participants, the literature was researched for simple tools to discriminate ID adults of MA 9 from

those above. We did this in spite of reports of 1- to 2- year operational delays between ID and non-ID subjects when matched for MA (Robinson & Robinson, 1976, as cited in Paour, 2001) that was not confirmed by Weisz & Yeates (1981) in their meta-analysis of 30 different studies covering 18 different conceptual domains.

### ***1.2.2.3. Additional Research for Cognitive Indicators for Children Aged 9 to 11***

Shaffer (1973) presented children aged 9 to 12 with a simple thought problem that seem to succeed in differentiating still-concrete-thinkers and those approaching or achieving formal operational thinking.

He asked them to consider the hypothetical proposition of being given a third eye and state, draw, and comment on such unreal scenario.

All the 9-year-olds placed the third eye on the forehead between their 2 natural eyes, and, possibly reflecting their still unimaginative stage of thinking, commented it as silly and boring. The elder group enjoyed the task and provides imaginative drawn site-varieties such as palm, or elevated from top of head.

### ***1.2.3. Quick Screening Tests Reviewed***

#### ***1.2.3.1. Quick IQ- substitute Tests***

Slosson's attempt in 1963 (*Slosson's Intelligence Test*, Hammil, 1968; Rotatori & Epstein, 1978) to devise a quick MH estimate for 4-16 years old for use by special education teachers, which, though popular in the 70s and 80s in the hands of

clinicians as an IQ-substitute test, was psychometrically weak and unrepresentative in sampling by modern standards (Kaufman & Lichtenberger, 2002).

Other more misguided early attempts included the combination test (Gioioso & Aderman, 1969) and 1986 Stanford- Binet abbreviated forms (Carvajal & Gerber, 1987) that failed on conceptual strength and purpose respectively. More recent ID “screeners” were also adaptive-behavior-omitting IQ/cognition quests designed as brief or lengthy cognitive competence tasks to test children subjects such as oddity or taxonomic tasks, cognitive contrast, tasks tapping memory or rhyming skills, with some or little apparatus set-up in a testing environment and mostly using a instructional format; primarily intended to differentiate against learning disabilities, achieving mixed results (Bornholt et al, 2004; Fletcher et al, 1999, Fletcher et al, 2004; Scott et al, 1993; Scott et al, 1997; Scott et al, 1998; Scott et al, 2000).

#### ***1.2.3.2. Quick Cognitive Functioning Test: Mini-Mental State Examination***

On the other hand, a very good and comprehensive global index of cognitive function for clinicians’ usage is the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) that has enjoyed exemplary psychometric status of reliability and concurrent validity (Folstein, Gregory, 1999). It has been used for diverse age-groups (Russel, Bornholt, & Ouvrier, 2002; Xu, Meyer, Huang, Du,

Chowdhury, & Quach, 2003) and patient samples (de Leon, 1993; Grace et al, 1995; Uhlmann, Larson, & Buchner, 1987; in Gregory, 1999) with translated and modified versions functioning world-wide (Teng & Chu, 1987 in Gregory, 1999). The original version takes 5 to 10 minutes to administer, and has the distinct feature of consisting of individually scored items that are very brief and mostly (16 out of 18) vocally-communicable, out of which only 2 require handy props as pencil, watch and a piece of paper (Folstein et al, 1975).

#### ***1.2.3.3. Quick Screening Tests for Non-professional Use***

Curiously, for non-professionals' usage, screening tests for ID adults are rare in the literature, local or abroad. This is vastly unlike the continual development of screening tests for other disabling functioning states such as Alzheimer's disease/dementia (Hopkins, Kilik, Day, Rows, & Hamilton, 2005; Lam, Chiu, Ng, Chan, Chan, & Li, 1998) and cognitive impairments in psychiatric disabilities (Allen, 1996; Chan, Ray, & Trudeau, 2001) which are available tools to health professionals not trained in specialized assessment techniques for obvious practical benefits (Hopkins et al, 2005).

The nearest parallel as such locally was a deliberate "helping police to identify persons with mental handicap" screening "street test" solution by the then Chief



Clinical Psychologist of the Social Welfare Department specifically for identification of ID persons by local police officers (Yau, 1994). It carried the distinct advantage of simplicity to conversationally interview a suspected unaccompanied ID person regarding his memory of his personal ID number and attendance in a special school or rehabilitation centre. The first criterion is of course invalid for its general assumption of poor memory. The second criterion no longer holds in the present local practice of integrative education. Hence, both are irrelevant.

#### ***1.2.3.4. Difficulties in Developing Quick Screening Tools***

The dismal amount of research efforts into screening tools for adult ID individuals may be due to a lot of factors as follows. The heterogeneous aetiological nature of mental deficiencies in different individuals in the way of organic (genetic, perinatal and postnatal) or cultural-familial (Dykens, 1998, in Burack, Hodapp, & Zigler, 1998) could be one; the high incidence and risk rates of associated psychopathology (anxiety disorder, depression), psychiatric problems (epilepsy), and substance abuse another (Luckasson et al, 2002). Moreover, the diagnostic role of adaptive behavior is constantly being redefined (Luckasson et al, 2002) and it is noted that different ID individuals have different trajectories of adaptive developments (Loveland & Tunali-Kotoski, 1998, in Burack, Hodapp, & Zigler, 1998).

Researchers that attend to content and construct validity of the multidimensional construct of ID, could still be waiting for its conceptual delineation of intelligence and adaptive behavior into distinct or aggregate behavior indicators to be operational for measurement (Luckasson et al, 2002). A brief test yielding inaccurate results could have labeling or social discrimination consequences, whereas development and refinement of standard-sized IQ tests and adaptive behavior tests already established upon adequate psychometric properties that are properly norm-referenced may be more fruitful.

#### ***1.2.4. Selection of Focus Group Domains***

A systematic comparison of scales and cognitive indicators reveals choices and overlaps. As a direct input to initiate, nourish but also to set parameters for Focus group discussion, a clear skeleton framework of possible intellectual and adaptive behavioral domains was prepared as shown in Appendix B.

##### ***1.2.4.1. Selection of Intellectual Functioning Domains as***

###### ***“Cognitive Functioning” Indicators***

Based on Inhelder’s model of operational diagnosis (1943, 1968), 3 Piagetian conservation tasks (which consists of task-types of number, length, and weight) and 2

seriation tasks were chosen (which consists of unidimensional and bi-dimensional task-types). Based on MMSE (Folstein, Folstein, & McHugh, 1975), short term memory, attention and calculation, language comprehension and expression were chosen. Based on Shaffer's hypothetical proposition (Shaffer, 1973), imagery was chosen (as third-eye-placement).

#### ***1.2.4.2. Selection of Adaptive Behavior Domains as***

##### ***“Daily Common Sense” Indicators***

Selective combinations of subsets of the 2002 AAMR model (Luckasson et al, 2002), the Vineland Adaptive Behavior Scale-Survey Form (“Vineland-S”, Sparrow, Balla, & Cicchetti, 1984); and the Home-based Training Package (HBTP) (Crawford & Tse, 1988) were made as summarized by appendix B.

The conceptual skills domain of the 2002 AAMR model has IQ elements as of language, reading and writing concepts that were replaced by the communication subset of the Vineland-S scale.

The Vineland-S scale has its 2 other subsets of socialization and daily living skills neatly corresponding to AAMR 2002's 2 remaining adaptive behavior domains as social skills and practical skills respectively.

For our local ID screening tool, we take these 2 domains together as its

“everyday common sense” component (of 3-components) to be matched against the locally developed Home-based Training Package (HBTP) (Crawford & Tse, 1988) which would have its 7 cores stripped partially, leaving out the 2 cores of motor skills and communication skills (Core I & V), to consist of 5 cores of self-care, domestic, community living, social interpersonal, and recreation skills.

### ***1.2.5 Criteria for Task Inclusion***

The finalized tasks needed to meet certain criteria:

- **Simplicity** - as examiners administering the tasks would not be specialists but layman with no training on psychological knowledge, the tasks need to be easy to understand and simple and quick to set up for absolutely observable verbal or gesture responses.
- **Clarity** - to fit the comprehension ability of the intellectually disabled examinee, the task-communicating language need to be direct, unambiguous and jargon-free; and in a suggestibility-free presenting style to safeguard against possible habitual yes/ no response sets.
- **Differentiating** - to fulfill the identification objective of not wanting to let pass false negatives, best items should have high difficulty values to allow only the very best of the intellectually disabled to give correct responses.

- Motivational - to overcome test anxiety and capture attention, tasks need to be interesting and presented in a playful rather than challenge manner, with warm well-done reassurances provided for examinees when eliciting right or wrong responses.

# *Chapter 2: Methodology*

## **2.1 Research design**

Employed here were a) Focus groups to achieve face validity of test items; b) personal semi-structured interviews with the target group to affirm content validity (relevancy) of, and to explore the ecological validity (specificity) of test items besides rapport establishment; and then c) standardized test administration to target (ID adults) and comparison groups (MA-matched non-disabled children subjects and matched-CA non-disabled adults), in the way of a quasi-experiment design.

## **2.2 Participants**

In the target group, participants were 68 intellectually disabled adults [2 of the 70 interviewed declined our scale test, which was much less than the expected one third drop-out ----- as the intended sample size was 50 to give sound effect size power (Nunnally, 1978)]. They were aged 25 to 53, 25 females and 42 males, all purposely sampled, through established contacts, for a good mix of different impairment levels, employment backgrounds, and social experience from the following: two Fu Hong Society sheltered work-shops (Aberdeen and Shek Wai Kwok branch), one St. James' Settlement sheltered workshop (Wanchai building), one

ID activity group (New Mileage Group) of the Joint Council of Parents of Mentally Handicapped People, and one self-advocacy group (Chosen Power). Although individual IQ scores or IQ impairment levels were not available on a reliable basis, rough mild/ moderate proportions as previously notified by corresponding management of these organizations that has allowed for a calculated guess of moderate/mild/border-line ratio of 4:3:1. As for the non-disabled matched-CA group, there were 50 adults aged 18 to 60 (taken as CA-matched to target ID adult group's 25 to 53), with 34 females and 16 males. A second non-disabled matched-MA group was of 50 children, aged 6 to 15 (MA-matched to target ID adult group's 4:3:1 ratio of CA-age levels of 6-8.9: 9-11.9: 12+), as 25 females and 25 males. Both comparison group participants were selectively recruited from attendants of a local dental clinic.

### ***2.3. Apparatus and Materials***

#### ***2.3.1. Apparatus & Materials for Focus Group discussion***

For Focus group meetings, meeting content recording utilities include recording machines and note-taking apparatus were made available, cognitive and behavioral discussion domains were as listed in Appendix B.

For *cognitive ability* assessment, we seek to adapt the Mini-Mental Status Examination (Folstein, Folstein & McHugh, 1975). For daily common sense

assessment, we seek to adapt the two domains of *daily living skills and socialization* of the Vineland Adaptive Behavior Scales to the relevant 5 skill cores listed in the locally formatted Home-based Training Package for Persons with Mental Handicap (Crawford & Tse, 1988). For *general knowledge*, we depend wholly on the focus group members to produce a consensus of the scope and depth of such pertaining to our regional, cultural and social context.

### ***2.3.2. Apparatus & Materials for Target Group Participant Interviews***

For interviews with intellectually disabled participants, interview topic-questions and trial tasks were listed in appendixes C.

### ***2.3.3. Apparatus & Materials for Scale Administration***

For actual test administration, props were set up to correspond to each task-item as follows with local bank-notes and HKSAR government-produced metal coins used where applicable (See appendix F):

Item 1 ---- 2 rows of 5 one dollar coins, all with the numeric-1 facing upwards, were glued by blue-tag onto the centre of a tri-folded inward-folding 100-gram white paper, with distance between coins as zero-spaced in one row and 20mm apart in the other.



Item 2 --- 2 rows of 5 assorted coins, all with numeric side facing upwards, glued by blue-tag onto centre of a tri-folded inward-folding 100-gram white paper, with distance between coins as zero-spaced in one row and 20mm apart in the other.

The assorted coins were in the same serial order of one 2-dollar coin first, then one 1-dollar coin, then one 50-cents coins, two 20-cents coins last in both rows.

Item 3 --- Two pieces of clean A4 papers were rolled separately inwards on themselves 4 times and hard-pressed neatly to the same longitudinal form and size.

Item 4 --- Two 20-dollars bills were rolled separately inwards on themselves 4 times and hard-pressed neatly to the same longitudinal form and size.

Item 5 and 6 --- One foreign-make mechanical family-use weighing platform.

Item 7 --- Four assorted coins, as 5-dollar, 1-dollar, 50-cents, and 10-cents.

(All coins align to the same unidimensional increase in surface and thickness---- this item probes unidimensional seriation).

Item 8 --- Five assorted coins, as 10-dollar, 5-dollar, 1-dollar, 50-cents, and 10-cents. (The 10-dollar coin is smaller and yet thicker than the 1-dollar coin. This item probes bi-dimensionality of surface area and thickness).

Item 9, 10, and 11 --- One piece of note-card, of the size 6 x 4 inches.

(Item 12 was intended to probe the adaptive behavior of self-budgeting).

Item 13 --- Two cans of unopened soft-drinks: one diet-coke, one

conventional.

Item 12 being the only interview item, needs no props but with participants' responses plus all other item responses graded by assistant researcher on the same note-card recording participants' own written and drawn responses to Item 9, 10, and 11, but on the opposite side (See appendix F).

## ***2.4. Procedures***

### ***2.4.1. Procedures for Focus Group meetings***

Two separate focus groups, each consisting of 8 to 9 invited speakers (invitation letter please see Appendix A), in the way of parents and caregivers of and service-providers for intellectually disabled people, special education teachers, and applied social studies academics but with no intellectually disabled adults, were run in a conference room in a special education school in Tai Hang Tung, Kowloon early January, 2007, with specific cognitive and behavioral discussion domains as listed in appendix A to search for local intellectual disabled adults' characteristic cognitive and adaptive behavior patterns. Efforts were made to manage an even share of view-expression time for all invited guests amongst between-guest interactions.

### ***2.4.2. Procedures for Interviews with Target Group Participants***

Following that, separate get-to-know you open dialogue person-to-person semi-structured interviews covering topic-questions and tasks listed in appendix C were carried out in January and February, 2007 by one same researcher-interviewer (who would become the sole researcher-examiner for test administration later to eliminate inter-rater discrepancies), assisted by an manual recording researcher with all of the initially chosen target sample of 70 intellectually disabled adults, each lasting for 8 to 20 minutes, in rooms provided by the participating workshops and societies. Personal or caregiver consent for participation was obtained previously by personal, written, or phone notification through workshop managers and activity leaders communicating with participants and their family members. Some workshop adults were encouraged to join as practice for later job interviews. Most parents accepted to help answer the research's aim. Each participant was thanked for their time and provided with small gifts in the way of tooth-paste.

Thirteen task-items were established as relevant and standardized for props set-up and accompanying verbal instruction to be ready for administration.

#### ***2.4.3. Procedures for Scale Administration***

All participants were administered the 13-item test on a one-to-one basis using standardized briefing wordings as shown by Appendices D and E.

Each target group participant attending work or activity was previously notified by his or her work or activity leader of the re-visit of the researcher team (since their last chit-chat general interview). All were informed of “new games” that the researchers were to bring for them to try out, thanked and asked at the end of the testing session which item they individually like most.

Scale administration for the invited non-disabled participants were the same except site of administration was a dental clinic surgery-room (with granted co-operation and facilities from concerned dentist), run after their appointed dental work.

[The briefing wording accompanying each item-task was designed to avoid response sets (and guessing) ----- the correct answer being neither one mentioned].

Post-test debriefing was carried out for all the participants to tell them the study objective.

## **2.5. *Data analysis***

SPSS.12 was employed for data analysis.

## *Chapter 3: Results and Findings*

### **3.1 Item Analysis**

The current research aims to develop a scale with proper construct validity. It intends to measure latent psychological constructs of adult ID targets through the difference in behavioral performance scores between them and other ability groups. In order to do that, sound statistical properties have to be established for individual items and their inter-relatedness to give the scale its internal consistency. The optimal strengths of their item statistics are listed as follows (Kaplan & Saccuzzo, 1989; Nunnally, 1978).

a) Item Discrimination (D value) is an item's ability to reflect a participant's ability level or total score, called its corrected Item-Total Correlation. The corrected implies that the total is not the sum of all item scores, but the sum of item scores without including the item in question. A good D-valued item separates high scorers from low scorers and high item discrimination is desirable. Consideration would be given to drop items with discriminations lower than .30; so D should be  $>.30$  and positive values closer to 1 are most desirable.

b) Item difficulty ( $p$ -value) reflects the proportion of people who answer the item correctly. Item is easy for high  $p$ -values; item is hard for low  $p$ -values. If  $p$ -value

= 1.0, everyone does task correctly, and there would be no variability in item scores.

Our intention to use the scale to achieve efficient identification for including (i.e. failing) the least mildly ID implies low  $p$ -values for most items is desirable, i.e., a mixture of fairly difficult ones is the rule here.

c) Cronbach's Alpha is a measure of the internal consistency of our scale that ranges from 0 to 1 with higher values indicating higher reliability. It should be  $> .60$

The results of Item Difficulty ( $p$ ) and Item Discrimination ( $D$ ) are presented in Table 1 and Table 2 respectively.

Table 1

Item Statistics of item difficulty ( $p$ ) for ID participants' total score means sorted from hardest (lowest  $p$ -values) to easiest (highest  $p$ -values)

	<i>Mean</i>	<i>S.D.</i>	N
Item 12	.06	.24	68
Item 6	.07	.26	68
Item 5	.09	.29	68
Item 4	.12	.33	68
Item 3	.12	.33	68
Item 2	.22	.42	68
Item 1	.22	.42	68
Item 11	.27	.45	68
Item 8	.28	.45	68
Item 10	.37	.49	68
Item 9	.53	.50	68
Item 13	.54	.50	68
Item 7	.78	.42	68

Note. Mean = Item difficulty  $p$ -value; e.g. for item 10, 37% of ID participants obtained the correct response to it. The  $p$ -value of item 10 is .37

Table 2 Item-Total Statistics in order of descending D-values to explore and discard low D-value items ( drop item if  $D < .30$ ; Cronbach's alpha  $> .60$ = good))

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correction	Cronbach's Alpha if Item Deleted
Item 3	3.54	8.61	.78	.84
Item 4	3.54	8.61	.78	.84
Item 1	3.44	8.19	.77	.84
Item 2	3.44	8.19	.77	.84
Item 5	3.57	8.85	.75	.84
Item 6	3.59	8.99	.72	.85
Item 12	3.60	9.23	.63	.85
Item 10	3.29	8.24	.62	.85
Item 8	3.38	8.54	.55	.85
Item 9	3.13	8.39	.54	.85
Item 7	2.88	9.21	.32	.87
Item 11	3.4	9.44	.21	.87
Item 13	3.12	9.78	.05	.89

Note. Corrected Item-Total Correction = Item Discrimination index, D value.

Scale's internal consistency, Cronbach's Alpha was, for 13 items, .86

Items 7, 11, and 13 were thus dropped for their unsatisfactory D-values. The item difficulty values of items 7 and 13, being the highest (.78 and .54) cause no debate. Item 11 has a fairly high (.27)  $p$ -value when compared to item 9 (.53) but the latter stays for its well above .30 (.54) D-value. The rest items show very high to moderately high  $p$ -values from .06 to .37

Results of item statistics of the thus shortened scale are presented in Table 3.

Table 3 Finalized 10-item Scale: Item-Total Statistics in order of descending D values

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correction	Cronbach's Alpha if Item Deleted
Item 3	1.96	6.58	.83	.89
Item 4	1.96	6.58	.83	.89
Item 1	1.85	6.25	.79	.89
Item 2	1.85	6.25	.79	.89
Item 5	1.98	6.82	.78	.90
Item 6	2.00	6.99	.73	.90
Item 12	2.01	7.18	.66	.91
Item 10	1.71	6.39	.59	.91
Item 8	1.79	6.52	.58	.91
Item 9	1.54	6.46	.53	.92

Note. Corrected Item-Total Correction = Item Discrimination index, D-values.

Scale's internal consistency, Cronbach's Alpha was, for 10 items, .91

### ***3.1.1 Split-half reliability checks***

Split reliability checks were run to secure against inappropriate splits that could jeopardize the scale's practical utility as shortened parallel forms. Results are presented in Table 4.



Table 4

Summary of Split-half coefficients and inter-form correlations of exploratory split-half combinations

Form A Items/ Form B Items combinations	Spearman-Brown Coefficient	Guttman Split-Half Coefficient	Correlation between Forms
A 1,3,5,10,12/ 2,4,6,8,9	.97	.96	.93
B 1,3,5,9,12/ 2,4,6,8,10	.96	.96	.93
C 1,3,5,8,9/ 2,4,6,10,12	.96	.96	.93
D 1,3,5,8,10/ 2,4,6,9,12	.96	.95	.92
E 1,3,5,8,12/ 2,4,6,9,10	.91	.91	.84
F 1,3,5,9,10/ 2,4,6,8,12	.91	.91	.84
G 1,2,3,4,5/ 6,8,9,10,12	.85	.85	.74

High values were achieved.

In detail, the data shows that Form-combinations A, B, C, and D are a small league apart from E and F and the rest, but greater split-half coefficients differences emerge between G and all the others. Examination of Form-pair contents reveals the very good split-half reliability for pairs that dissociate items 1, 3, 5 from the others but the best splits are combinations that group 1, 3, 5 together with either but not both of the pair 9 and 10.

Hence, 1, 3, 5, 9, plus any non-10 item or 1, 3, 5, 10, plus one “non-9” item would be the most favorable split format for one form.

### **3.1.2. Gender differences**

Though female ID adults obtained lower total score means ( $M = 1.32$ ,  $SD = 2.12$ ) than male ID adults ( $M = 2.57$ ,  $SD = 3.14$ ), the difference was not significant,  $t(65) = -1.95$ ,  $p = .056$ . Gender is thus concluded to have no significant effect on ID adult participants' total score means.

### **3.2 Across-group comparisons of scale item scores**

Independent t-test results are as follows:

ID adult participants obtained lower total score means ( $M = 2.07$ ,  $SD = 2.84$ ) than non-disabled adult participants ( $M = 9.82$ ,  $SD = .56$ ). The difference was significant,  $t(116) = 21.92$ ,  $p < .001$

ID adult participants obtained lower total score means ( $M = 2.07$ ,  $SD = 2.84$ ) than non-disabled children participants ( $M = 5.66$ ,  $SD = 2.54$ ). The difference was significant,  $t(116) = 7.09$ ,  $p < .001$

It could be concluded that there were statistical significant differences of total score means of our target ID adult group with, with other different-ability groups ---- both the ID adult group and non-disabled children group.

## *Chapter 4: Discussion*

### *4.1. Summary and Discussion of Findings*

#### *4.1.1. Scale Reliability and Item Statistics Findings*

Fulfilling established statistical criteria as advocated by Kaplan & Saccuzzo, (1989) and Nunnally (1978), the scale has produced a set of 10 task items with internal consistency greater than the desired value .60 (being .91). Item discrimination index was well above .30 (being .54 to .78), implying that all 10 items have satisfactory inter-item correlation and each is able to discriminate target participants. Split-reliability statistics show split-combination patterns achieving fair to very good values (.74 to .97) confirms a preference to split the items in a way to partition certain very closely inter-correlated item, such as item 1 and 2, 3 and 4, 5 and 6 (each item-pair is of exactly the same or very minimally different item-difficulty values as .22/.22, .12/.12, and .09/.07). This is coherent with the way such item pairs were designed in the first place that each such pair measurement respects conservation of the same physical quantity as number, length, and weight respectively. A less obvious necessity in split arrangement is to also dissociate items 9 and 10, which represents the same skill type measurement but of ascending difficulties ( $p$ -values as .37 and .53). Split-form combinations as A, B, C, D should suffice as optimal splits. Relatively

high item difficulty  $p$ -values were obtained for 9 items (.37 to .06) with the remaining item-9 (10-8 subtraction) at .54. This is in concordance with our aim to establish hard-to-pass items to identify the cream of the mildly ID adult citizens.

#### ***4.1.2. Scale Performance across Target and Comparison Groups***

It could be concluded that there were statistical significant differences of between-groups total score means, that increased from ID adult group to non-disabled children group to non-disabled adult group. The essence of such ranked difference lies in the ability of our developed scale to screen and identify ID adults on the basis of their typical task performance as statistically proven to be drastically unlike that of the CA-matched non-disabled adult participants, and also not as good as the MA-matched non-disabled children.

#### ***4.1.3. Gender Bias***

Statistical analysis shows non-significant scale score mean difference for the 26 female and 42 male ID adult participants, thus provides reassurance against gender bias. The scale thus developed could be used on both sexes.

#### ***4.1.4. Practical Consideration for Individual Item Functionality***

For efficient whole scale screening for ID status, item difficulty values are first

considerations, so the item sequence 12, 6, 5, 4, 3, 2, 1, 8, 10, 9 (in descending item “hardness”) should be preferred. However, items should, for practical utility feasibility, be considered for their functionality needs as regards the following factors: set up speed and precision, inherent motivationality versus test anxiety or test awareness, verbal duration and semantic content, and ethical concerns. We would now examine items for individual usage. Item 12 possess the highest  $p$ -value but carries with it terrible privacy-encroaching dangers and therefore ethical concerns. As it was designed to probe self-budgeting abilities, and being a interview question (the only one), it is not as robust as others task-items in avoiding lied answers or attracting demeaning feelings. It was included in the first place due to valuable experience-sharing from caregivers of ID adults in the Focus Group that did find ecological substantiation during subsequent rapport-building interviews carried out. We propose refrigerating thus item unless usage is absolutely appropriate and necessary.

Items 1 to 8 are considered to enjoy a motivating character as Piagetian tests (Paour, 2001) while item 13 (the coke choice), though discarded as a credible test item for its low discrimination D-value, may find re-birth as a initial rapport-building ice-breaking interpersonal contact-initiation gesture or simplest dialoguing when revised as “hey, have a coke, choose one”. Item 9 and 10 seem to achieve

unmistakable talk-transference for its almost talk-free set-up, yet, besides item difficulty considerations, they carry relatively higher risks of test awareness and anxiety on top of possible stereotyping dangers which Piagetian items (items 1 to 8) possibly elicit least. Of these, item 8 seems most motivational as it involves total concentration and actual subject's physical handling of shown props that bears less test awareness than item 9, which, with its highest *p*-value and utter simplicity may find real-life significance in being transformed (when asked for verbal response instead of subject putting it in writing) into a very quick and efficient identifying tool for the minimally verbal and much-disabled, if used with gentle manner and sensitivity. If the subject makes a correct response, the same could be tried out for item 10 as an immediately following conversational item.

In everyday situations, as made used of by a layperson to test his ID examinee, items have to be robust across real here-and-now test-situations e.g. item 3 and 4 are faster and easier to set up than items 1 or two as the former pair have minimal requirement for props material ----- 2 same-valued banknotes rather than 10 same-valued or 4 or 5 pairs of different valued coin-pairs. Verbal length is similar for items 3 and 1, but item 3 is inherently more motivational as it initially requests examinee gestured involvement to confirm same length of the 2 banknotes. However, in practical application, items 5 and 6, their heavy semantic content aside, may suffer

from props set-up obstacle; and when props not being available, the hypothetical imagination required may hinder proper reasoning process when a true-life testee may just give up and resort to guessing.

To sum up, items 3, 4, 8 could be best items for simplicity and differentiating.

## ***4.2. Implications***

### ***4.2.1. Theoretical Implications***

The cross-group comparisons have shown that our target ID adult group participants perform least well than both other-ability groups. This is in line with most study findings reviewed except one. Our results seem to confirm the delay as reported by Robinson & Robinson (1976, as cited in Paour, 2001) between his ID subjects and MA-matched non-ID subjects. However, we do not consider our total scale score data confirming or disapproving such, as different tasks were used in their and our studies (we used a mixture of Piagetan and non-Piagetian task-items). [The very fact that, even if we had over-matched to include older children (teenagers), and yet a statistically significant whole-scale performance difference is obtained between thus older children and some of our younger-looking ID adults only serve more to our scale's benefit as a quick screening tool to screen for ID adults].

To sum, our data, which were based on operationality score usage in fabricating most

of its items, seem to provide, at least, partial support to Inhelder's ID reasoning level principles.

#### ***4.2.2. Implications for Our Research Objective***

Statistical analyses and practical functionality considerations seem to point to our scale as an effective and efficient quick screening tool to provide fast psychological rather than physical identification of our ID fellow citizens. To fulfill such wish, the still unknown factors influencing the hypothetical common man as an examiner need to be explored. We would now set out the absolute (yet theoretical) minimal requirements that we should ask of this layman person in that he or she is of the intellect level and experience to be able to firmly distinguish another fellow Cantonese-speaking person as to whether this latter person is an adult or non-adult.

The tool should help disciplinary officers, rehabilitative personnel, caregivers, activity instructors, work supervisors and all those of the willing and welcoming section of the general public to be more welcoming to their old or new ID adult acquaintances, mates or clients, through a deeper realization of their cognitive and adaptive disabilities.

#### ***4.2.3. Implications for Possible Misuse:***



Because these are Piagetian structural indicators, and as tests to check operational levels (e.g. conservation being the land-mark achievement of concrete operation (Piaget & Inhelder, 1974)), their incorporation into our scale could enable a direct revelation of our ID adult subject's operational stage status that may promote understanding and acceptance of their reasoning strengths or weaknesses. Our identification objective could be enriched in such a way but it needs stressing that association of observed behavioral responses (correct/wrong answering) with developmental stages is not the primary objective in this project and further extrapolation to diagnostic purposes may weaken our tool for its primary screening usage and may foul it with labeling contempt.

Inhelder herself repeatedly emphasizes that operational diagnosis should only be part of and not the sole element in a whole clinical examination (on top of the fact that she did not realize the necessary consideration of social expectations) (Paour, 2001). We must stress that no matter how impressive or even eye-opening these tests could be when eliciting wrong answers, they are as such designed for easy set-up usage by a layperson who may have little or no understanding of the psychological principles nor clinical knowledge influencing such behavioral responses.

#### ***4.3. Methodological Limitations***

The researcher failed to gain access to individual IQ impairment status of the target participants, making IQ-correlation of our results impossible. Accurate IQ scores or mild/ moderate grading would have allowed for external criterion validity check of our data which would be even more statistically persuasive than our reliance on comparing data with other ability groups. Our Focus Group did not generate enough feasible ideas due to its heavy proportion of non-clinical personnel.

Our pre-scale-test semi-structured interview questioning has provided rich qualitative data that eventually find little relevant usage for our research objective other than establishing rapport. Such data loss to be set against the time-consuming nature of thus clumsy interviewing methodology calls for deserved criticism as it had caused delayed scheduling that has resulted in good chance missed of follow-up re-test data collection that would have enable fuller statistical test-re-test reliability analysis. All in all, this research suffers from the shortcomings of a convoluted design and its over-ambitious implementation. Across-group *t*-test comparisons for individual items were not attempted due to the time factor.

#### ***4.4. Cross-cultural Concerns***

The obvious question is whether this new and yet psychometrically tested tool could have results generalizable outside the research-test contexts. This depends on a)

sameness of apparatus, b) personal qualities of examiner and examinee, and c) their interaction. As yet, we have no empirical proof of serious complications arising from b); while c) would certainly yield to interpersonal chemistry but if task set-up was properly replicated, the screening robosity (correct answer not provided) inherent in most, if not all the items, would achieve positive screening. For sameness of apparatus, it would be the sameness of rolled-up lengths inherent in any same numerical-valued bank-notes or mass-manufactured typing paper (e.g. A4-sized paper) anywhere, in whatever countries or regions. Coins may need re-adaptation but it should be easy to find assorted coins in size increments of more than one dimension.

We find secure reliance on Piagetian conservation and seriation tasks thus Inhelderian-validated as to bear universal stages of intellectual development to be applicable across-race or skin-color.

Caution, however, is due for items 9 and 10. These two could suffer from cultural practices that place heavy educational or job-related training upon simple calculations that could see their item difficulty values dissolved by repeated practice and becoming non-functional, their learnt component should be made aware of. The rule of thumb should be, if in doubt, always test subjects on the whole scale (with special consideration for including item 12). That was the way it obtained its significant results in this study, i.e., as a set score.

#### ***4.5. Suggestions for Future Research***

We suggest better planning and proper supported requests for valid IQ-status records for ID participants prior to testing targets in research involving ID subjects. The methodological design should include interviews with clinical psychologists, and other experienced rehabilitation personnel to provide personal insights about specific cognitive and adaptive behaviors.

We also see a true necessity to explore layman qualities through controlled true trials of screening for willing real-examinees.

#### ***4.6. Conclusion***

Quick identification of ID people has always been a necessity locally. Immense sadness still clouds certain families for evaporation of their loved ones. Psychology should always try to help turn up strategic tools and knowledge to harmonize contacts with ID people as a integration-willing minority group or as a wandering soul missing the direction of home and family.

## References

- Allen, C. K. (1996). *Allen Cognitive Level Screen 1996: Test Manual*. Colchester, CT: S&S.
- American Psychiatric Association (APA) (2000) *Diagnostic and Statistical Manual of Mental Disorders*, 4<sup>th</sup> edn. American Psychiatric Association, Washington, DC.
- Bornholt, L. J., Spencer, F. H., Fisher, I. H., & Ouvrier, R. A. (2004). Cognitive Screening for Young Children: Development and Diversity in Learning Contexts. *Journal of Child Neurology*, 19, 5,313-317.
- Carvajal, H., & Geber, J. (1987). 1986 Stanford-Binet abbreviated forms. *Psychological Reports*, 61, 1, 285-286.
- Clarke, A. M. & Clarke, A. D. B. (1974). Prospects for prevention and amelioration of mental subnormality: an overview. In A. M. Clarke & A. D. B. Clarke (Eds.), *Readings from Mental Deficiency: The Changing Outlook*. Methuen & Co.: London.
- Chan, C. L. F., Ray, S., & Trudeau, S. (2001). The development of the Cantonese version of the Cognitive Level Screen. *Occupational Therapy International*, 8, 4, 287-300.
- Chosen Power. ( 2002). *Advocacy New Era: Chosen Power in Ten Years 1992- 2002*.

Hong Kong: Chosen Power, People First Hong Kong

Crawford, N., & Tse, J. (1988). *Home-based Training Package for Persons With Mental Handicap*. Hong Kong: Social Welfare Department.

Drew, C. J. & Hardman, M. L. (2007). *Intellectual Disabilities Across the Lifespan*. (9<sup>th</sup> ed.). Pearson: Upper Saddle River, New Jersey.

Dykens, E. M. (1998). Maladaptive behavior and dual diagnosis in persons with genetic syndromes. In J.A. Burack, R.M. Hodapp, & E. Zigler, (Eds.), *Handbook of Mental Retardation and Development* (pp.542-562). New York: Cambridge University Press.

Equal Opportunities Commission. (2001). *Equal Opportunity Commission Report on procedural and training needs for Immigration Department's handling of Disabled Persons*. Hong Kong: Equal Opportunities Commission.

Fletcher, K. L., Blair, C., Scott, M. S., & Bolger, K. E. (2004). Specific Patterns of Cognitive Abilities in Young Children with Mild Mental Retardation. *Education and Training in Development Disabilities*, 39, 3, 270-278.

Fletcher, K. L., Scott, M. S., Deuel, L. S. & Jean-Francois, B. (1999). A comparison of the cognitive capabilities of young children with mild mental retardation or learning disabilities. In T. E. Scruggs, & M. A. Mastropieri (Eds.), *Advances in Learning and Behavioral Disabilities*. (Vol. 13, pp. 65-93). Greenwich: JAI Press Inc.

Folstein, M. F., Folstein, S.E., & McHugh, P.R. (1975). 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Folstein, M. F. (1998). Mini-mental and son. *International Journal of Geriatric Psychiatry, 13*, 290-294.

Gregory, R. J. (1999). Foundations of Intellectual Assessment: The WAIS-III and Other Tests in Clinical Practice. Boston: Allyn & Bacon.

Gioioso, J. V., & Aderman, M. (1969). The combination test as a quick screening device to differentiate levels of retardation. *Psychological Reports, 25*, 3, 843-848.

Government Secretariat, Health and Welfare Bureau. (1999). *Hong Kong Rehabilitation Program Plan (1998-99 to 2002-2003): Towards a New Rehabilitation Era*. Hong Kong: Hong Kong Government Printer.

Hammil, D. (1968.) The Slosson Intelligence Test as a quick estimate of mental ability. *Journal of School Psychology, 7*, 4, 33-37.

Hodapp, R. M., & Dykens, E. M. (1994). Mental retardation's two cultures of behavioral research. *American Journal on Mental Retardation, 98*, 675-687.

Hogg, J., & Langa, A. (2005). In J. Hogg, & A. Langa (Eds.), *Assessing adults with*

*Intellectual Disabilities: A Service Providers' Guide*. Malden, MA: Blackwell Publishing.

Hopkins, R.W., Kilik, L. A., Day, D. J. A., Roes, C. P., & Hamilton, P. (2005). The Brief Kingston Standardized Cognitive Assessment- Revised. *International Journal of Geriatric Psychiatry*, 20, 3, 227-231.

Hu, F. Y. (2007, March 21). Down's syndrome pair walking tall. Post office workare thriving after landing jobs. *South China Morning Post*, Section C, p. 4.

Intellectually disabled man missing for 30 hours found. (February, 2007). *Ming Pao*, Section A, p.6.

Kaplan, R. M. & Saccuzzo, D. P. (1989). *Psychological Testing: Principles, Applications, and Issues*. (2<sup>nd</sup> ed.). Pacific Grove, California: Brooks/Cole Publishing Company.

Katz, N. (1992). *Cognitive Rehabilitation: Models for Intervention in Occupational Therapy*. Butterworth-Heinemann: Stoneham, M A.

Kaufman, A. S. & Lichtenberger, E. O. (2002). *Assessing Adolescent and Adult Intelligence* (2<sup>nd</sup> ed.). Boston: Allyn & Bacon.

Lam, L. C. W., Chiu, H. F. K., Ng, K. O., Chan, C., Chan, W. F., & Li, M. (1998). Clock-face drawing, reading and setting tests in the screening of Chinese elderly adults. *Journal of Gerontology*, 53B, 6, 353-357.



- Lee, C. K. (1997). Community Rehabilitation: New Milestones for Rehabilitation Services. In Y. L. Choi, & T. W. Yeung (Eds.), *Mental Handicap Service for Adults in Hong Kong: Reviews and looking Forward*. Hong Kong: Chinese Bookshop.
- Loveland, K. A., & Tunali-Kotoski, B. (1998). Development of adaptive behavior in persons with mental retardation. In J.A. Burack, R.M. Hodapp, & E. Zigler, (Eds.), *Handbook of Mental Retardation and Development* (pp.521-541). New York: Cambridge University Press.
- Luckason, R., Coutler, D. L., Polloway, E.A., Reiss, S., Schalock, R. L., Snell, M.E., Spitalnik, D. M., & Stark, J. A. (2002). *Mental retardation: Definition, classification, and systems of support* (10<sup>th</sup> ed.). Washington, DC: American Association on Mental Retardation.
- New City Plaza Fire causing intellectual disabled man went missing. (May, 2007). *Ming Pao*, Section A, p. 1.
- Ngai, L. K. H. (1997). Target-focused Supported Employment Services. In Y. L. Choi, & T. W. Yeung (Eds.), *Mental Handicap Service for Adults in Hong Kong: Reviews and looking Forward*. Hong Kong: Chinese Bookshop.
- Nunnally, J. C. (1978). *Psychometric Theory* (2<sup>nd</sup> ed). New York: McGraw-Hill Book Co.

- Paour, J-L. (2001). From Structural Diagnosis to Functional Diagnosis of Reasoning: A Dynamic Conception of Mental Retardation. In A. Tryphon, & J. Voneche (Eds.), *Working with Piaget: Essays in Honour of Barbel Inhelder*. Psychology Press: London.
- Pasnak, R., Wilson-Quayle, A., & Whitten, J. (1998). Mild retardation, academic achievement, and Piagetian or psychometric tests of reasoning. *Journal of Developmental and Physical Disabilities, 10*, 23-33.
- Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities: Conservation and atomism* (A.J. Pomerans, Trans.). London: Routledge & Kegan Paul. (Original work published 1941)
- Rotatori, A. F., & Epstein, M. H. (1978). The Slosson Intelligence Test as a quick screening test of mental ability with profoundly and severely retarded children. *Psychological reports, 42*, 3, 1117-1118.
- Russell, L., Bornholt, L., & Ouvrier, R. (2002). Brief cognitive screening and self concepts for children with low intellectual functioning. *British Journal of Clinical Psychology, 41*, 93-104.
- Scott, M. S., Fletcher, K. L., Jean-Francois, B., Urbano, R. C., & Sanchez, M. (1998). New Screening Tests to identify Young Children at risk for mild learning problems. *Journal of Psychoeducational Assesement, 16*, 302-314.
- Scott, M. S., Fletcher, K. L., & Martell, B. (2000). Selecting components for a

screening test to identify three-year-olds at risk for mild learning problems. *Education and Training in Mental Retardation and Developmental Disabilities*, 35, 208-221.

Scott, M. S., Perou, R., Claussen, A. H., & Lois-Lynn, S. (1997). A cognitive portrait of grade school students with mild mental retardation. *International Review of research in mental retardation*, 20, 157-189.

Scott, M. S., Perou, R., Greenfield, D. B., & Swanson, L. (1993). Rhyming skills: Differentiating among mildly mentally retarded, learning disabled, and normally achieving students. *Learning Disabilities Research & Practice*, 8, 4, 215-222.

Shaffer, D. R. (1973). *Children's responses to a hypothetical proposition*. Unpublished manuscript, Kent State University.

Social Welfare Department. (2005). *Directive on Prevention and Handling of Lost/Missing of Intellectually Disabled persons*. Social Welfare Department: Hong Kong S.A.R.

Sparrow, S., Balla, D.A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales: Interview Edition/ Survey Form Manual*. Circle Pines, MN: American Guidance Service.

Tong, C. K., Chan, Y. L., & Yau, T. Y. (2002). *Pleasure in Communicating: Rehabilitative integration Project*. (in Chinese). Hong Kong Family Welfare Society: Hong Kong.

- Tsang, O. (2007, May 2). Splashing out. *South China Morning Post*, Section C, p. 3.
- World Health Organization. (2001). *International Classification of Functioning, Disabilities, and Health (ICF)*. Geneva: Author.
- Weiaz, J., & Yeates, K. (1981). Cognitive development in retarded and non-retarded persons: Piagetian tests of the similar structure hypothesis. *Psychological Bulletin*, 90, 151-178.
- Weisz, J., & Zigler, E. (1979). Cognitive development in retarded and non-retarded persons: Piagetian tests of the similar sequence hypothesis. *Psychological Bulletin*, 86, 831-851.
- Weiten, W. (2004). *Psychology: Themes and Variations*. (6<sup>th</sup> Ed.). Wadsworth/Thomson: Belmont, CA.
- Xu, G., Meyer, J. S., Huang, Y., Chowdhury, M., & Quach, M. (2003). Adapting Mini-Mental State Examination for dementia screening among illiterate or minimally educated elderly Chinese. *International Journal of Geriatric Psychiatry*, 18, 609-616.
- Yau, P. (1994). *Helping Police to identify persons with mental handicap*. Hong Kong: Social Welfare Department.

## *Appendix A*

### **Letter of Invitation to prospective Focus Group candidates**

你好，我是劉殷良，城市大學後學士文憑心理科(P/T) (postgraduate diploma in psychology, P/T,晚間上課)的學生，因選擇跟隨一位復康/智障科的專門學者, Dr. John TSE (謝永齡副教授)做一個學術 project/dissertation 研究，現唐突邀請閣下參與一個主題會會議 (Focus group meeting)。

Dr. John TSE 致力發展一套有可靠性和準繩度 (reliable and valid) 的量表，簡潔快速地去初步甄別香港社會中的成年了的輕微智障人士(Mildly intellectual disabled / Mildly mental handicapped)。

我希望能邀請閣下作為對本地此類人士的認識和心得作出支持，促成其事。

此乃一個很小規模的研究，但可成為日後有力、成熟研究的基礎或開始。我們的研究水準有賴 Focus Group 各成員的經歷和熱心的交流。

Dr. John TSE 指示我希望能尋找到與智障人士於成長、教育、訓練等方面有深入了解和接觸的社會中各階層，角色人士，所以本會包括智障人士和其家長、特殊教育工作者，及家居照顧/協助人士；學者方面：包括社工、復康科、心理學科專家，一共約 10 位一起會議座談。

Dr. John TSE 提醒我務必請閣下作為受邀者在不設先限下提供任何意見，我會在開會前，耳聽手錄或經 e-mail 和任何方便閣下的聯絡途徑收集好，並能於第一次會議開始先作粗略範疇的藍本。

以下是可供刺激思維的範疇骨架的例子，如智商、生活/自顧能力、常識/教育程度、言語表達/溝通特徵等等。我們期望能參考閣下的寶貴掌握方法和思維，令我們能初步集合為數大約 30 項特徵作為 30 條量表的項目(Scale items)。交予一群成年智障人士和其他控制比較組別，例如：(成年非智障人士、12 歲非智障人士或 9 至 12 歲人士)，試驗後而製成有統計學術量度的最後 8 至 10 條項目。

我們請求閣下出席一共兩次的Focus group meeting會議，地點暫定為九龍南山邨。現附上Dr. John TSE所提供的兩頁「測試弱智快速工具」以作參考。請提供寶貴時間及意見，我們希望能在1月(1月20至26日)舉行第一次會議。

劉殷良 謹啓

Dr. John TSE @2677 7865

edward lau @e-mail: [edward2tooth@yahoo.com.hk](mailto:edward2tooth@yahoo.com.hk)

@mobile: 9703 6525 @office: 2392 4499

p.s. 此 project 第一階段的 Literature Review / Proposal Draft 可供 e-mail 發送參閱提供意見。

## *Appendix B*

### **Domains for Focus Group Discussion**

#### **Adaptive Behavior domains :**

**[daily common sense] :**

- **social skills/ socialization, practical skills/ daily living skills** (self-care, domestic, community living, social interpersonal, occupational and recreation skills).
  - ----- AAMR model (Luckasson et al, 2002);
  - Vineland Adaptive Behavior Scale-Survey Form (Sparrow, Balla, & Cicchetti, 1984); Home-based Training Package (HBTP) (Crawford & Tse, 1988) -----

**[general knowledge] : at the discretion of Focus Group.**

#### **Intellectual Behavior domains :**

**[Cognitive functioning] :**

- **Short term memory, attention and calculation, language comprehension and expression.**
  - ---- MMSE (Folstein, Folstein, & McHugh, 1975) -----
- **Piagetian conservation and seriation tasks**
  - ---- Inhelderian reasoning indicators (Inhelder, 1943,1968) ----
- **Imaginery**
  - ---- Shaffer's hypothetical "the third eye" proposition (Shaffer, 1973)----

## *Appendix C*

### **Semi-structured Interview Question Topics and Tasks**

#### **(Based on Focus Group feedbacks)**

[Hello, I am XXX, please come in and chat, please sit down]

1. Do you like chatting? (me too)
2. How old are you?
3. Let's see how good your memory is, what did you have for breakfast this morning?
4. Was it delicious?
5. Do you go eat at McDonald's often?
6. **Let's try saying "Uncle MacDonald" in Cantonese backwards, try.**
7. Do you want to write it down first?
8. If you buy a MacBurger for 8 dollars, and you give the man 10 dollars, how much would you get back as change?
9. If you buy a whole Macdonald meal for 24 dollars, and you give the man a 50 dollar bill, how much would you get back as change?
10. How would you react if someone in the street collide into you or step on your foot?
11. What would you do if you collide into someone in the street? Who taught you that?
12. If someone ask you to change his 2 dollar coin, what would you do?
13. If he ask you to change his 2 dollar coin into 10 cents, what would you do?
14. You know all the local Hong Kong coins?
15. **Try this, which row of coins have more money (total amount)?**
16. **Could you tell these Hong Kong banknotes from one another, try tell me.**
17. What cartoon books do you like reading? Why?
18. What TV cartoons do you like watching? Why?
19. Do you like watching "Police News"? Why?
20. If you found yourself loss on the street, what would you do?
21. To whom would you ask for help, a police officer or ordinary guy?
22. To whom would you ask for help, a man or a woman?
23. If you were lost and you met me, if I lend you my mobile, who would you call? Why?
24. Do you know your home's telephone number?
25. Do you know your HK Identification Card number?
26. Do you have good friends? Good workmates?

27. What do you think of the differences between very good friends and ordinary friends?
28. What do you think of men painting their finger-nails?
29. What do you think of men wearing ear-rings?
30. What do you think of men playing with dolls?
31. Do you read newspapers? Which parts?
32. Do you go to Library often? What do you borrow from the library?
33. Are you happy working here? What do you like doing here most? Why?
34. What did you do before joining here?
35. Where was your school? Were you happy there? Why?
36. Have you heard about people cutting their wrists? What do you think of them?
37. Would you do that to yourself? Why (not)?
38. If your boss here tells you he will give you a raise of 2 thousand dollars on the condition that you go to school again for 2 weeks to study and take notes and exams, all you need is to pass the exams, would you go? Why?
39. How do you see yourself? (Good? Hard-working? Lazy?)
40. If I say this: my English is "limited company", what does it mean?
41. Have you heard of "ear soft"? If I say that I am "ear-soft", what does it mean?
42. Do you know what "black Friday" means?
43. Can you tell me what "value system" means?
44. Are you a humorous person? What is humor?
45. If you put on 10 lbs of weight in 2 weeks, what would you do? Would you go seek help from slimming companies? Why (not)?
46. Do you mind if others cheat you?
47. If the person who cheats you is your mother, what do you do?
48. If you have to lend me one of these 3 things, which would you lend me? I just use them for 2 minutes when you are watching me here, your tooth-brush, you pajama, or your mobile? Why (the chosen item)?

[Very good, I will bring a weighing machine and other coins to play games with you next time, thank you very much.]



## *Appendix D*

### **Item Briefing: English version**

1. Piagetian number conservation item

This is 1 row of coins, this is another row, which row has more money? Not which row longer/shorter, it asks which row has more money. ([this row]/ pointed) well done.

2. Piagetian number conservation item

This is 1 row of coins, this is another row, which row has more money? Not which row longer/shorter, it asks which row has more money. ([this row]/ pointed) well done.

3. Piagetian length conservation item

These 2 strips of paper are of same length, use your hands to check, same length, OK? ([yes]) Now, it changes (examiner pulling the 2 strips away from one another horizontally until only about half overlapping) which is longer? Which is actually longer? ([this]/ pointed) well done.

4. Piagetian length conservation item

These 2 strips of 20 dollar notes are of same length, use your hands to check, same length, OK? ([yes]) Now, it changes (examiner pulling the 2 strips away from one another horizontally until only about half overlapping) which is longer? Which is actually longer? ([this]/ pointed) well done.

5. Piagetian weight conservation item

Look at this weigh, you try standing on it like me, try, OK, please sit down, so I stand like this, if I raise 1 foot to stand like this, would I be lighter or heavier? ([lighter/ heavier]) well done.

6. Piagetian weight conservation item

Look at this weigh, you try standing on it like me, try, OK, please sit down, so I stand like this, if I bend my knees and squat like this, would I be lighter or heavier? ([lighter/ heavier]) well done.

7. Piagetian unidimensional seriation item

Here are 4 coins, please arrange them by their circular sizes in a row, from small to big. (done) well done.

8. Piagetian bidimensional seriation item

Here are 5 coins, please arrange them by their circular sizes in a row, from small to big. (done) well done.

9. Subtraction item

What is  $10-8$  equal to? Please write down the whole calculation sequence, take your time ([down in a column or as a row?]) Up to you, take your time. (done) well done.

10. Subtraction item

What is  $50-24$  equal to? Please write down the whole calculation sequence, take your time ([down in a column or as a row?]) Up to you, take your time. (done) well done.

11. Third-eye item

Here, let's draw, you have 2 eyes, if a fairy gives you one extra eye, i.e. to give you a third eye, and you can place it anywhere on your body, where would you put it? all up to you. (drawn) well done.

12. Pocket-money item

The money you use to buy soft drinks or stuff to eat, how often does your family person give to you?

13. Soft drink Coke item

Here are 2 cans of coke, one is little sugar, people say it's healthier, one is ordinary more sugar, people say it makes you fat, which would you choose? ([this]/pointed) well done.

## *Appendix E*

### **Item briefing :Cantonese (Colloquial) version**

#### 1 皮亞傑數量常衡題

呢度有一行銀仔，呢度又有另外一行銀仔，边行多啲錢呢？唔係边行長啲、短啲，係边行多啲錢嘢？（「呢行」/指咗） 好，答得好好。

#### 2 皮亞傑數量常衡題

呢度有一行銀仔，呢度又有另外一行銀仔，边行多啲錢呢？唔係边行長啲、短啲，係边行多啲錢嘢？（「呢行」/指咗） 好，答得好好。

#### 3 皮亞傑長度常衡題

呢兩條紙，係一樣長嘅，你自己用手度吓，係唔係一樣長先？（「係」） 嘩，依家咁，（問者左右拉開至一半左右長短） 边條長啲呢？實在係边條長啲呢？（「呢行」/指咗） 好，答得好好。

#### 4 皮亞傑長度常衡題

呢兩條 20 蚊銀紙，係一樣長嘅，你自己用手度吓，係唔係一樣長先？（「係」） 嘩，依家咁，（問者左右拉開至一半左右長短） 边條長啲呢？實在係边條長啲呢？（「呢行」/指咗） 好，答得好好。

#### 5 皮亞傑重量常衡題

你睇吓呢度有個磅，你好似我咁企上去，試吓玩吓嘅，好，你請坐低，嘩我本來咁企嘅，若果依家我縮起一隻腳，咁單腳企，我會輕咗定重咗呢？（「輕咗/重咗」） 好，答得好好。

#### 6 皮亞傑重量常衡題

你睇吓呢度有個磅，你好似我咁企上去，試吓玩吓嘅，好，你請坐低，嘩我本來咁企嘅，若果依家我踏低縮埋個身，我會輕咗定重咗哩？（「輕咗/重咗」） 好，答得好好。

#### 7 皮亞傑跟類題

嘩呢度四個銀仔，請你將佢地跟圓形嘅大細，由細排到大俾我睇吓。（做咗） 好，做得好好。

### 8 皮亞傑跟類題

喺呢度五個銀仔，請你將佢地跟圓形嘅大細，由細排到大俾我睇吓。(做咗) 好，做得好好.

### 9 數學題

10-8 等如幾多呢？請你寫出嚟，要成條式連答案咁寫出嚟，慢慢寫（「直式或打橫寫呀？」）任得你鐘意，慢慢寫。(做咗) 好，做得好好.

### 10 數學題

50-24 等如幾多呢？請你寫出嚟，要成條式連答案咁寫出嚟，慢慢寫（「直式或打橫寫呀？」）任得你鐘意，慢慢寫。(做咗) 好，做得好好.

### 11 三眼題

喺，玩吓畫公仔先，依家你有兩隻眼，若果神仙佢俾多一隻眼你，即是俾第三隻眼俾你，任得你將第三隻眼放係你身體上任何一個位置，你會放係邊呢？任得你。(畫咗) 好，畫得好好.

### 12 零用錢題

你買汽水飲，買嘢食嘅錢，你屋企人幾耐俾一次你用架？

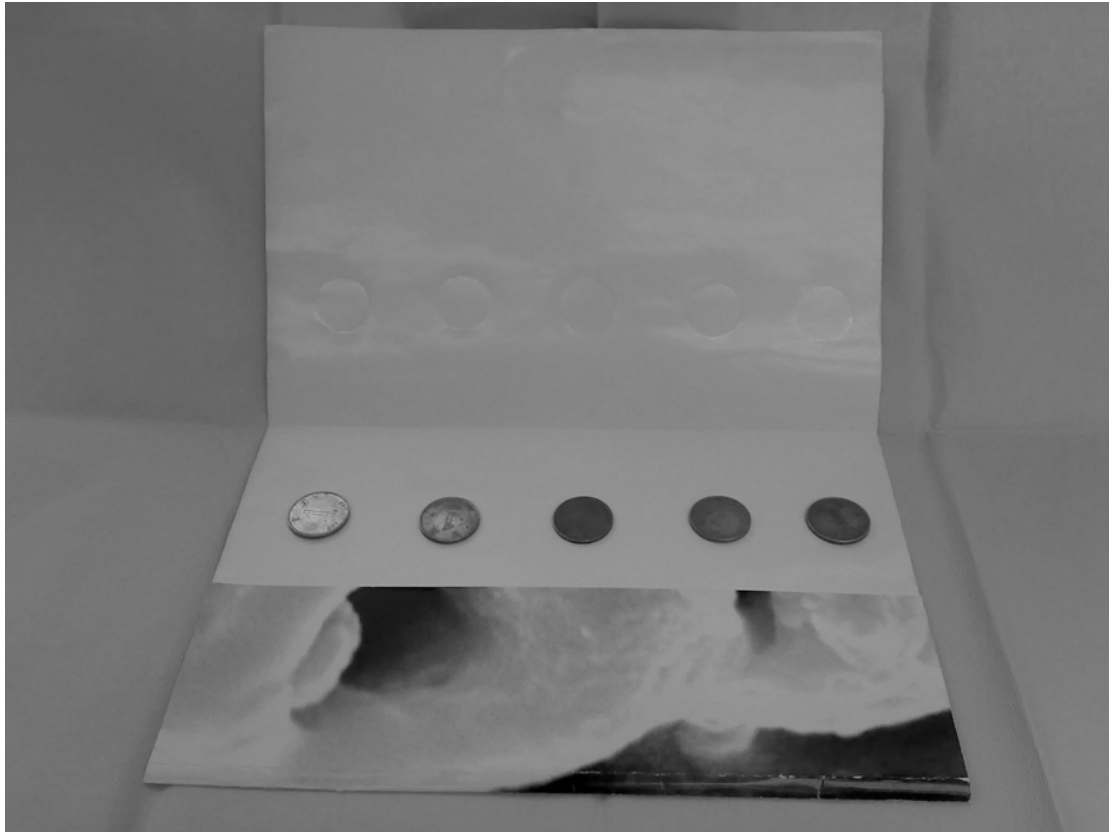
### 13 汽水可樂題

呢度兩罐可樂嘅，一罐少糖，啲人話健康啲，一罐正常多糖，啲人話會飲咗會肥啲，若果任你揀，你會揀邊罐？（「呢罐」/指咗）

*Appendix F*

Photographs of task items

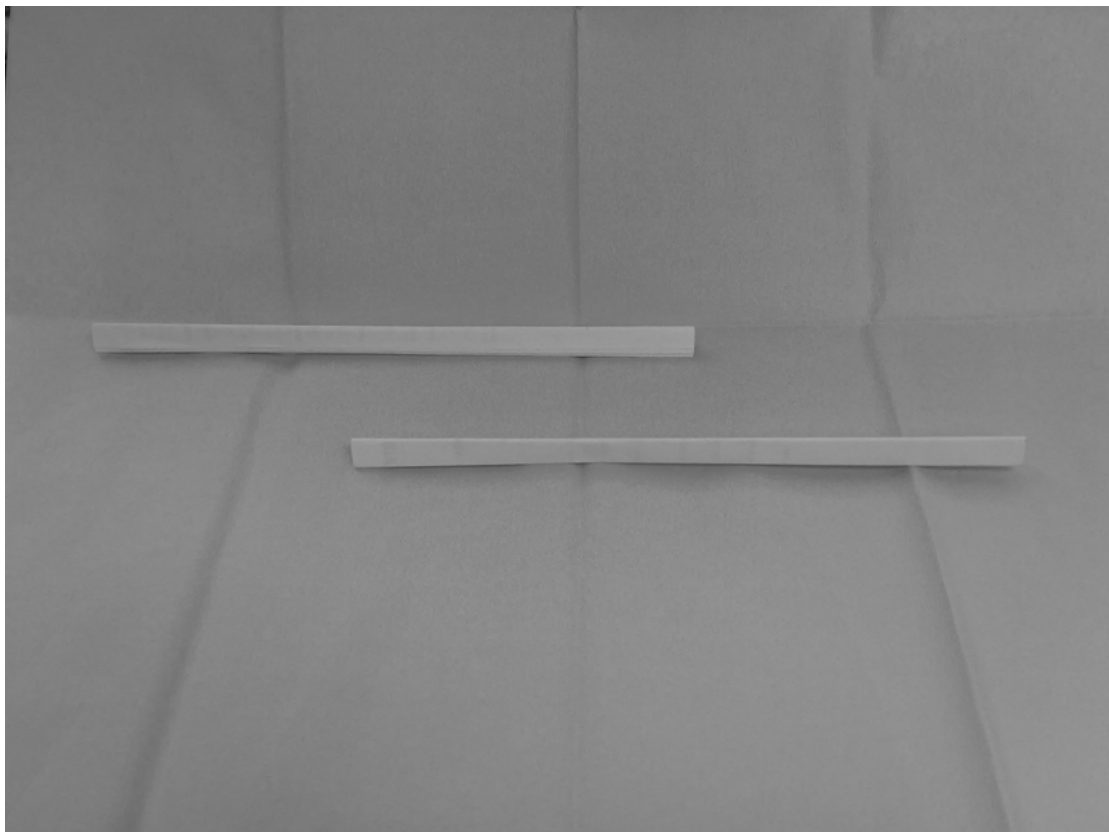
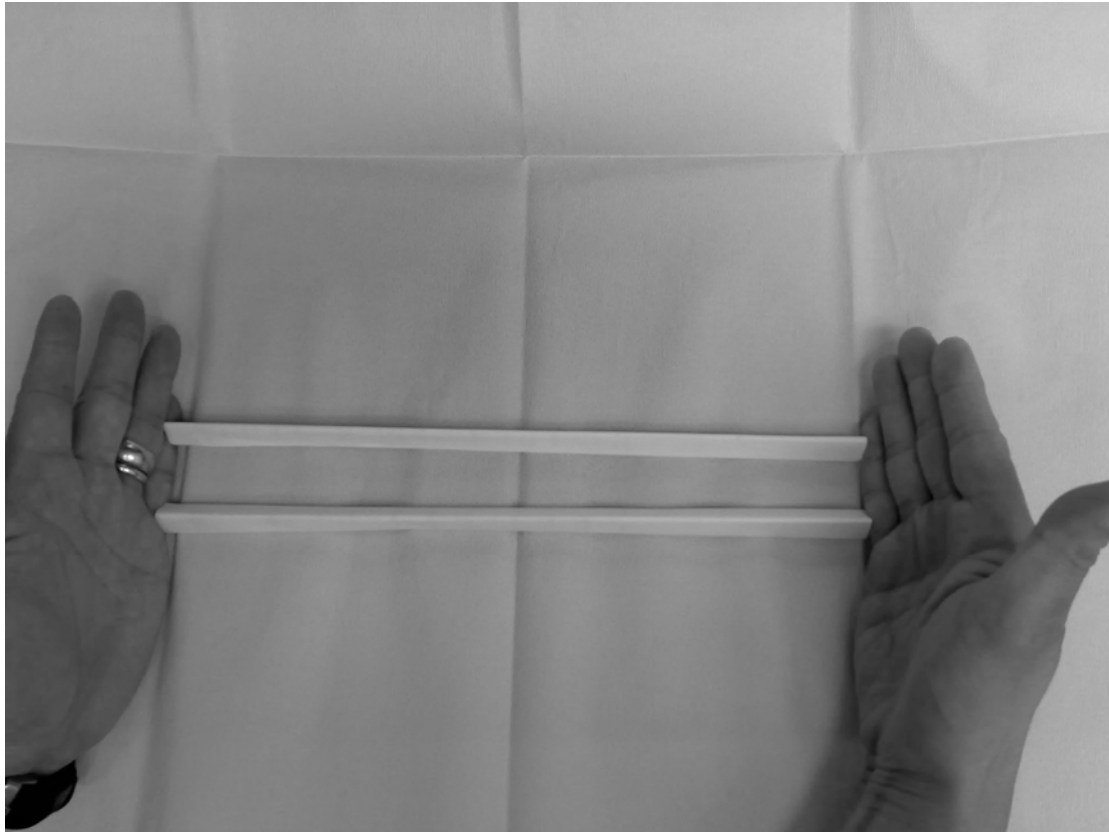
Item 1



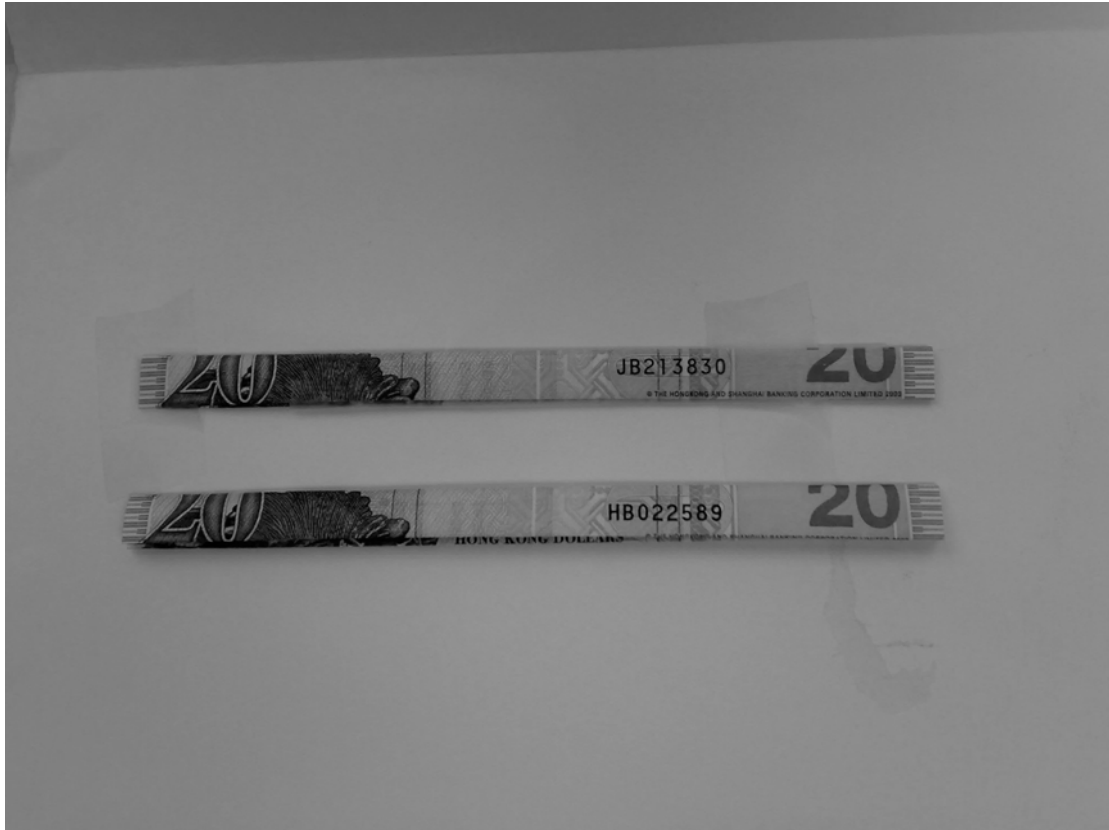
Item 2



Item 3



Item 4





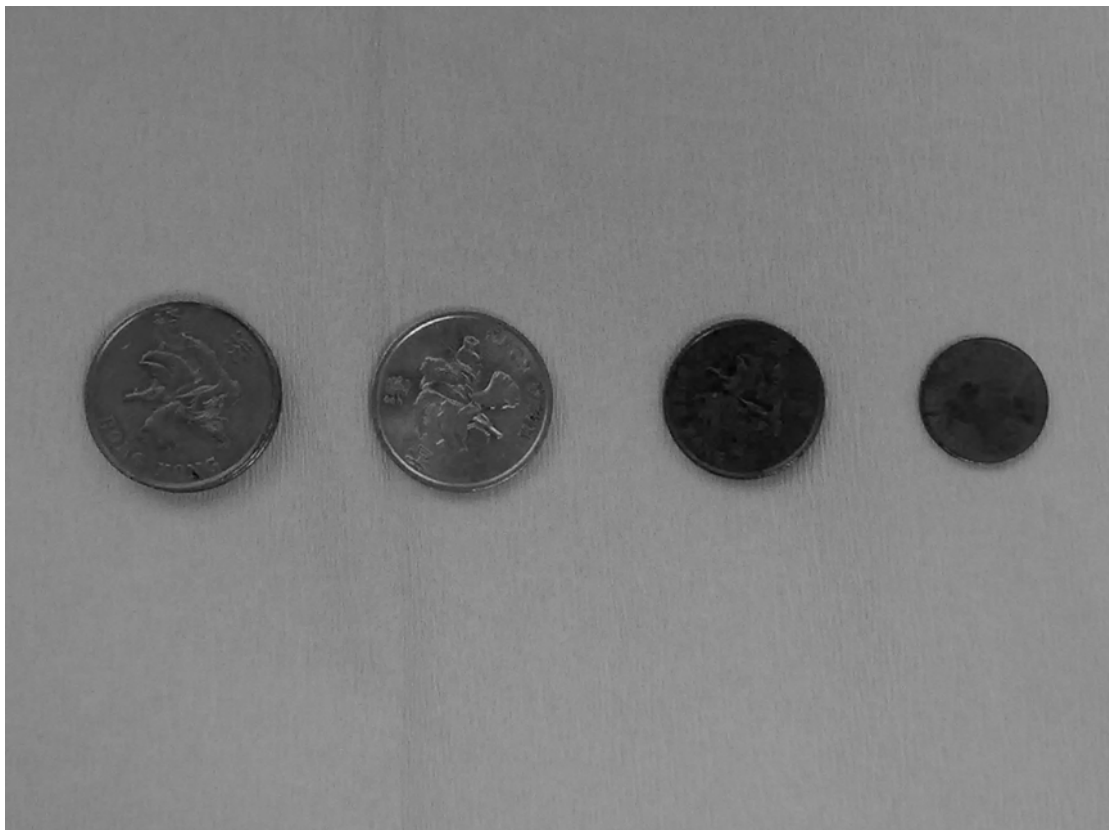
Items 5 & 6



Item 7



(Above: coins as presented; Below: Correct answer)



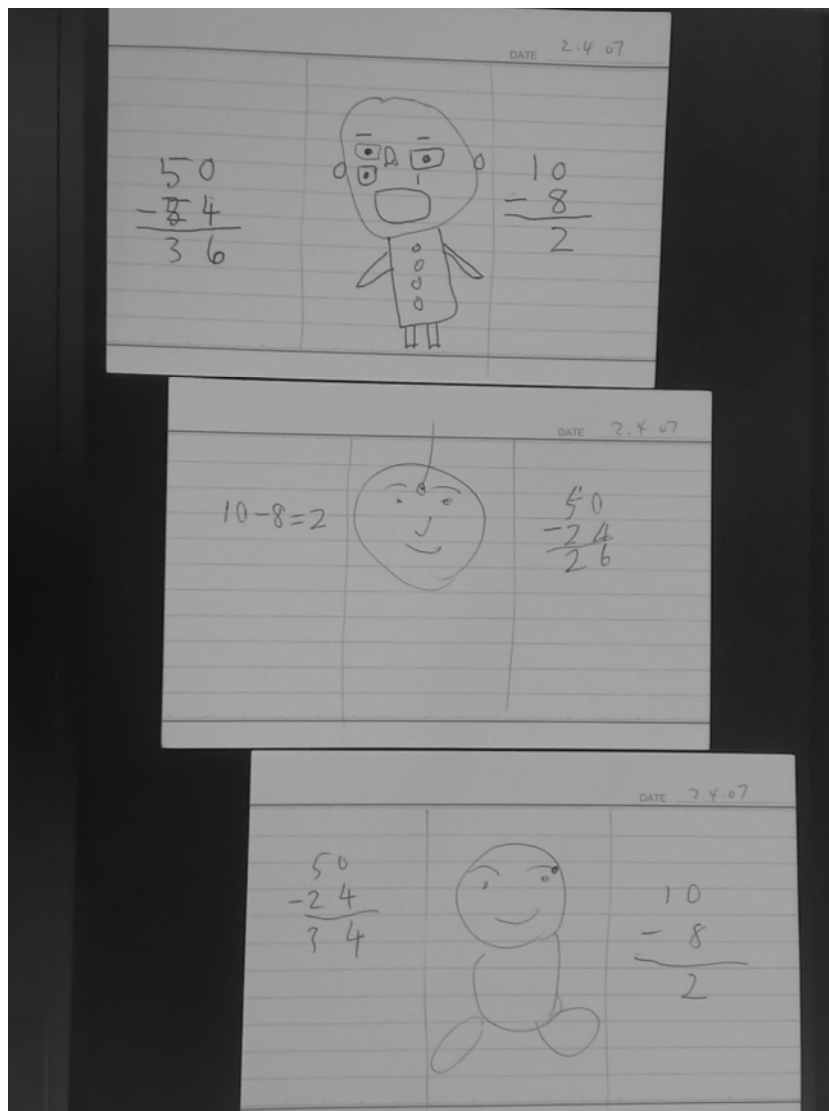
Item 8



(Above: coins as presented; Below: correct answer)



Items 9, 10, 11



Item 13

