THE VALIDITY STUDY OF THE TEST OF NONVERBAL INTELLIGENCE, THIRD EDITION: TONI-3

WORAYA SROYTHONG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE (CLINICAL PSYCHOLOGY) FACULTY OF GRADUATE STUDIES MAHIDOL UNIVERSITY 2008

COPYRIGHT OF MAHIDOL UNIVERSITY
Thesis
Entitled

THE VALIDITY STUDY OF THE TEST OF NONVERBAL INTELLIGENCE, THIRD EDITION: TONI – 3

Ms. Woraya Sroythong
Candidate

Assoc. Prof. Sudsabuy Chulakdabba,
M.D., Dip (Psychiatry)
Major-Advisor

Assoc. Prof. Chirdsak Kowasint,
Ed.D. (Research and Curriculum Development)
Co-Advisor

Prof. Banchong Mahaisavariya,
M.D.
Dean
Faculty of Graduate Studies

Assoc. Prof. Sucheera Phattharayuttawat,
Ph.D. (Psychology)
Chair
Master of Science Programme in Clinical Psychology
Faculty of Medicine, Siriraj Hospital

Copyright by Mahidol University
Thesis
Entitled

THE VALIDITY STUDY OF THE TEST OF NONVERBAL INTELLIGENCE, THIRD EDITION: TONI-3

was submitted to the Faculty of Graduate Studies, Mahidol University for the degree of Master of Science (Clinical Psychology) on
July 14, 2008

Ms. Woraya Sroythong
Candidate

Chair

Assoc. Prof. Sudsabuy Chulakdabba, M.D., Dip (Psychiatry)
Member

Asst. Prof. Raviwan Nivataphand, M.D.
Member

Assoc. Prof. Chirdsak Kowasint, Ed.D. (Research and Curriculum Development)
Member

Prof. Banchong Mahaisavariya, M.D.
Dean
Faculty of Graduate Studies
Mahidol University

Dean
Faculty of Medicine, Siriraj Hospital
Mahidol University
ACKNOWLEDGEMENTS

This thesis could not be succeeded at all without assistance from several people. I would like to thank Associate Professor M.D. Sudsabuy Chulakadabba the major-advisor and Associate Professor Dr. Chirdsak Kowasint the co-advisor and Associate Professor Kanokrat Sukhatunga the chairman, who always gave useful advice and helped in review and correction this thesis. These people also gave encouragement researcher along the way of studying.

Thank you to the director and all staffs of Wat Arunamarintaram for kindly helps and cooperation in data collecting.

Moreover, this thesis could not be completed if without a supporting scholarship from the Siriraj Graduate Scholarship, which allowed the research to be a success.

Lastly, I would like to thank of my family and all my dearest friends for their help, love, care and cheerful encouragement. Thank you so much.

Woraya Sroythong
THE VALIDITY STUDY OF THE TEST OF NONVERBAL INTELLIGENCE, THIRD EDITION: TONI-3

WORAYA SROYTHONG 4937418 SICP/M

M.Sc. (CLINICAL PSYCHOLOGY)

THESIS ADVISORS: SUDSABUY CHULAKADABBA, M.D.,
Dip (PSYCHIATRY), CHIRDSAK KOWASINT, Ed.D.
(RESEARCH AND CURRICULUM DEVELOPMENT)

ABSTRACT

The main objective of this research was to study the characteristics of measurement of TONI-3 test in terms of criterion-related validity, construct validity and practicality. Through random sampling, the study sample consisted of 380 primary school children from Wat Amarin Tharam School, Thailand.

Results of the study showed that the TONI-3 test had criterion related validity to CPM test at a moderately high level (r=0.488) and had construct validity that was more capable to categorize intelligence into extreme groups than CPM (Colored Progressive Matrices Test). Regarding practicality, the comparison between group and individual testing indicated that TONI-3 had more effectiveness with individual testing. Therefore, in practicality, TONI-3 was more subordinate than CPM because TONI-3 was only able to test individually. Besides, the study of comparison of composite scores as classified based on gender and age showed that different genders indicated similar levels of intelligence with average scores having the tendency to increase when age was increased.

It was found that TONI-3 was not appropriate for use by an examiner who had never been trained before because TONI-3 had some certain conditions for practice although TONI-3 could be considered as a screening test that the examiner could use to supplement standardized intelligence testing instruments or to confirm the results of other tests.

KEY WORDS: VALIDITY / TEST OF NONVERBAL INTELLIGENCE / NONVERBAL INTELLIGENCE / INDIVIDUAL TESTING

60 pp.
การศึกษาความตรงของแบบทดสอบ TEST OF NONVERBAL INTELLIGENCE , THIRD EDITION : TONI – 3 (THE VALIDITY STUDY OF THE TEST OF NONVERBAL INTELLIGENCE , THIRD EDITION : TONI-3 )

วรญา สร้อยทอง 4937418 SICP/M

วท.ม. (จิตวิทยาคลินิก)

คณะกรรมการควบคุมวิทยานิพนธ์: สุดสบาย จุลกทัพพะ, พ.บ., อ.ร. (จิตวิทยาคลินิก)
เข็มศักดิ์ โมราชัย, ก.ศ.ด. (การวิจัยและพัฒนาหลักสูตร)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่ศึกษาและวิเคราะห์ของแบบทดสอบ TONI – 3 ในด้านความตรง (Validity) โดยพิจารณาจาก ความตรงตามเกณฑ์สัมพันธ์ (Criterion – Related Validity) ความตรงตามโครงสร้าง (Construct Validity) และการปฏิบัติได้จริงของแบบทดสอบ (Practicality) กลุ่มตัวอย่างได้มาจาก การวิจัยการสุ่มอย่างง่าย โดยศึกษานักเรียนชั้นประถมศึกษาปีที่ 3 จากโรงเรียนวัดอมรินทราราม

ผลการวิจัยพบว่ามีความตรงตามเกณฑ์สัมพันธ์ โดยมีความสัมพันธ์ในระดับปานกลาง (r =.488) กับแบบทดสอบ CPM และมีความตรงตามโครงสร้างโดยการจับกึ่งระดับความสามารถเชาวน์ปัญญาในระดับ Extreme ได้ดีกว่าการจับกึ่งระดับแบบทดสอบ CPM ด้านในด้าน Practicality ผลของการเปรียบเทียบ ด้านการทดสอบแบบรายกลุ่มและรายเดี่ยวพบว่าแบบทดสอบ TONI-3 จะมีประสิทธิภาพดีเมื่อด้านการทดสอบเป็นรายบุคคล แต่ละวิ่งในด้าน Practicality มีความต้องการแบบทดสอบ CPM เพราะท่าการทดสอบแบบทดสอบ TONI-3 ได้เฉพาะรายบุคคล นอกจากนี้ในการศึกษาเปรียบเทียบค่าคะแนนเฉลี่ยในด้านร่างกายและสุขภาพพบว่าเพศที่แตกต่างกันมีระดับคะแนนไม่แตกต่างกันแต่เมื่อมีค่าเฉลี่ยของคะแนนในแบบทดสอบ TONI – 3 จะสูงขึ้นเมื่ออายุเพิ่มขึ้น

ในการนำแบบทดสอบ TONI-3 ไปใช้ในการคัดกรองเชาวน์ปัญญา พบว่าแบบทดสอบ TONI-3 ไม่เหมาะสมที่จะนำไปใช้คัดกรองเชาวน์ปัญญาโดยที่ไม่ได้ด้านการฝึกอบรม ทั้งนี้เพราะแบบทดสอบ TONI – 3 มีเงื่อนไขในการปฏิบัติ อย่างไรก็ตามแบบทดสอบ TONI – 3 เป็นเพียงแบบทดสอบคัดกรองทางเชาวน์ ปัญญา ดังนั้นควรจะใช้ร่วมกับแบบทดสอบทางเชาวน์ปัญญามาตรฐานอื่นๆ เพื่อยืนยันผลการตรวจร่วมกัน

60 หน้า
# CONTENTS

<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ENGLISH ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>THAI ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURE</td>
<td>ix</td>
</tr>
</tbody>
</table>

## CHAPTER

### I  INTRODUCTION

- Background and Significance of the Study 1
- Objectives 4
- Research Instruments 4
- Benefits 5
- Definition of Terms 5

### II  LITERATURE REVIEW

- Intelligence 7
- Properties of Psychology Tests 12
- The Non – Language Intelligence 16
- Conceptual Framework 29

### III  MATERIALS AND METHODOLOGY

- Population and Sample 30
- Research Instruments 32
- Data Collection 33
- Data Analysis 35
CONTENTS (Continued)

<table>
<thead>
<tr>
<th>IV</th>
<th>RESULTS AND DISCUSSION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Results</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V</th>
<th>CONCLUSION AND RECOMMENDATIONS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conclusion</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Recommendations</td>
<td>54</td>
</tr>
</tbody>
</table>

| BIBLIOGRAPHY | 56 |
| BIOGRAPHY    | 60 |
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 The Classification of standard score by the TONI-3 test</td>
<td>19</td>
</tr>
<tr>
<td>2 Classification of the CPM test</td>
<td>26</td>
</tr>
<tr>
<td>3 Number of sample group as distributed in gender and age</td>
<td>32</td>
</tr>
<tr>
<td>4 Classification of intelligence (IQ) as calculated from composite scores of TONI-3 and CPM group tests</td>
<td>37</td>
</tr>
<tr>
<td>5 Result of the analysis of fitness in the classification of intellect of the TONI-3 in comparison with CPM in a group test</td>
<td>38</td>
</tr>
<tr>
<td>6 Classification of IQ level as calculated from composite scores of TONI-3 and CPM tests in both individual and group tests using 60 school children with ages ranging from 9 years to 9 years and 11 months</td>
<td>39</td>
</tr>
<tr>
<td>7 Results of the analysis of difference of IQ value from TONI-3 and CPM as classified according to group and individual testing</td>
<td>41</td>
</tr>
<tr>
<td>8 Results of the analysis of the score difference between TONI-3 and CPM based on Composite Scores as classified according to gender</td>
<td>42</td>
</tr>
<tr>
<td>9 Results of the analysis of the difference in scores from TONI-3 as classified according to age</td>
<td>43</td>
</tr>
<tr>
<td>10 Results of the analysis of the difference in scores in age pairs at different age levels from TONI-3 test</td>
<td>44</td>
</tr>
</tbody>
</table>
## LIST OF FIGURE

| Conceptual framework of the study | 29 |
CHAPTER I
INTRODUCTION

Background and Significance of the Study

Studies of school children have found that children have many problems that affect the education of these children in their normal grade level particularly those problems related to intelligence which normally occur among children with higher or lower than normal intelligence level. In the case when intelligence is higher than the normal level, the child is characteristically smarter than his age, thinks with creativity and maybe has a special skill in one aspect such as music or arts although this child still remains as a child who has other needs just like other normal children. Children belonging to this group tend to become easily bored with school at their normal grade level because learning at this level may not offer anything interesting to motivate them to become excited to learn more. For a child with above normal intelligence, most of these problems come as a result of the people around him who do not understand the nature of the child belonging to this group and not able to facilitate appropriately the child’s needs and capabilities thus resulting problems relate to his ability to adapt to a situation such as separating himself from the group of his friends, feeling bored from studying things other than what he himself is interested with or restricted doubts which are promoted using the child’s intellectual abilities. Yet there is a lack of mood response according to a child’s age (1). Meanwhile, a child with below normal intelligence can never learn anything in school, cannot keep up with his classmates or is only able to study things particular to that specific grade level but with significant difficulties in the early grade level which contain subjects which were studied more and lessons were more difficult (2). Thus measuring the intelligence of a school child is important in sorting out children who have higher or lower than normal intelligence in order to promote or develop appropriately the level of a child’s intelligence.
The measurement of intelligence has several methods with many kinds of tools, in which measuring intelligence, individual testing or group testing may be applied. One of the most popular methods of testing intelligence in Thailand and in many other countries is the Stanford-Binet Intelligence Scale (3), an individual testing of intelligence which can be applied to small children. This test which was originated by Binet and Simon in 1905, was then called the Binet-Simon Scale and consisted of questions about decision-making, logic and understanding. The test was improved continuously and the first time was in 1908 into a new test where questions were divided into sections based on the age level of those being tested to create a standard measurement of mental age. The second improvement was in 1911 when it was used for children from 3-18 years of age and scoring was more detailed. Another popular testing method was called the Wechsler Intelligence Scale which was created in 1939 by Wechsler. Initially, the test was focused as an intelligence test for adults and was improved in 1955 to become the Wechsler Adult Intelligence Scales (WAIS) and was used as a method of testing persons from 16 to 75 years of age. In 1945, an intelligence test was created for children and was called the Wechsler Intelligence Scales for Children (WISC). Likewise, this intelligence test was later improved in 1974 to become the WISC-R and was used for children aged 6 to 16 years. Then in 1991, it was again improved and was called the WISC III for children of same range of age. Both Stanford Binet Intelligence Scale and Wechsler Intelligence Scale were applied as verbal scale intelligence tests.

For Test of Nonverbal Intelligence (TONI), it was started in 1982 by Brown, Sherbenou and Johnsen and was used as an individual test for persons from 6 to 89 years and 11 months of age. It was later improved in 1990 and was called TONI-2. Most recent improvement was in 1997, becoming TONI-3 and has been used to sort out a child’s intelligence especially in cases where children are less intelligent and could not be tested by the verbal test method. This test can be used for individuals who have problems in hearing, speaking, reading and writing, and for those with brain abnormalities and learning problems including those with abnormalities in perception, memory or semantics. The TONI-3 intelligence test was created in order to evaluate one’s sensual capacities based on general elements of intelligence (g-factor) according
to Spearman’s theory that considers the measurement of an individual’s capacity to observe, to perceive through sight, to think clearly and to use conflicting logic. Aside from this, TONI-3 is considered a test that is simple and uses a less number of tools including a shorter time to finish thus not allowing the child who is being tested to worry and at the same time provide assistance to the teacher to sort out children who might be prepared in terms of intelligence. Preparing children for the development of intelligence and sorting them out before and during primary education serve to benefit the teachers and parents of these children.

However, the most important thing about an intelligence test is its quality that could be seen from two important characteristics, namely: validity and reliability. Validity is the characteristic of being able to indicate whether the intelligence test can accurately measure what it wants to measure based on prior objectives. On the other hand, reliability indicates the correctness of the level of scores acquired from several measurements. Reliability has three important types: content validity, criterion-related validity and construct validity (5).

Realizing its importance, the researcher has therefore decided to study TONI-3 intelligence test based on the quality of testing and classification in comparison with CPM, a standard test that measures intelligence using its general elements. In a study by Suchira Pattharayuthawat, et al (6), CPM was used to test a group of Thai individuals, the results showed a high reliability of 0.91 when Kuder – Richardson 20 was used. From the study of Siri Udomphol (7), Benjawan Ransaewa (8), Sudarat Sirisakphanit (9), Janthanee Mungketklang (10), Wararat Inthapthim (11), Vichana Molichat 912) and Surirat Palakas (13), which used CPM to test Thais from all regions in the country, results showed a high reliability ranging from 0.90 to 0.92. Based on the results of these tests, the researcher has observed that CPM is an intelligence test that has a high reliability and which could serve as a standard in the comparison of reliability of nonverbal intelligence tests such as TONI-3.

As a consequence, the researcher became greatly interested to evaluate the validity of TONI-3 by studying the content-related validity and comparing it with CPM test and if validity is found it would be beneficial in one aspect as one type of intelligence test that is able to classify the intelligence level of Thai children or used
on a group of sick people or even children with communication problem in places such as in schools, hospitals or others in order to manage the education level or to identify guidelines for future health care.

**Objectives**

The primary aim of this research was to study the characteristics of TONI-3 intelligence test in terms of its validity, as follow:

1. **Criterion-Related Validity** by determining the relationship between the test scores from TONI-3 intelligence test and those from CPM intelligence test.
2. **Construct Validity** in terms of Age Differentiation to study the differences in the intelligence of school children of different ages.
3. **On the actual practicality** of the intelligence test based on the study of the benefits and limitations in the actual use of this intelligence test in terms of time period of implementing data collection and convenience in the use of the research tool.

**Scope of the Study**

**Population**

The population of this research consisted of both male and female school children with ages that ranged from 7 to 11 years and currently studying in 2007 from schools in Region 3 under the Committee on Basic Education of the Ministry of Education in Bangkok Metropolis.

**Research Instruments**

2. Colored Progressive Matrices (CPM), revised in 1956.
Benefits

1. Knowledge about the validity of TONI-3 when compared with the standard CPM intelligence test.

2. Knowledge about the limitations including the weakness and strength of using TONI-3 intelligence test to Thai children aged 7-11 years.

Definition of Specific Terms

Validity refers to the Criterion-Related Validity which identifies the relationship between TONI-3 and CPM, and Construct Validity in terms of Age Differentiation.

Practicality means the ability of actually using the intelligence test with consideration of its strength, weakness and limitations of the intelligence test in its usage.
CHAPTER II
LITERATURE REVIEW

In this study, the various literature and research reports related to this research had been divided into different sub-topics, as follow:

1. Intelligence
   Theoretical framework of intelligence
   Development theories of intelligence
2. Properties of research tools used for psychological tests
3. Non-language tests used for classifying intelligence test
   Non-Verbal Intelligence Test, Third Edition
   Colored Progressive Matrices Test
4. Conceptual framework of the study
Intelligence

For many years, numerous psychologists and educators have argued on the definition of intelligence and yet, nobody has made certain conclusions. Each one had a different meaning of intelligence based on his own reasoning and research. As a start, Binet (6) provided a definition of intelligence as a trend in using intelligence to understand something depending on the characteristics of that thing and can use his understanding and modify the solution, creativity in solving a problem, to achieve the objective and also as a characteristic of an ability to criticize oneself. Later Wechsler (16) stated that intelligence is the ability of an individual to adapt himself to the environment, ability to learn and solve problems together with the ability to think with abstraction, which was similar to Garret (15) who defined intelligence as composed of all abilities that are necessary to solve the problem based on one’s understanding and signs. On the other hand, Spearman (17) believed that intelligence has two components, namely: General factor (g-factor) which involves various behaviors such as logic or reasoning, problem solving and perception of all experiences, etc.; and Specific factor (s-factor) which is the ability for music, languages, etc.

To summarize, the meaning of intelligence is considered broad and covers many scopes, from the ability of an individual to learn and the ability to solve various problems that occur through the use of reasoning and the ability to understand varied issues together with the understanding of many signs and symbols, thinking with abstractions and the ability to adapt to a new environment without any past related experience. Based on those definitions, the psychologists who created both TONI-3 and CPM intelligence tests were in direct agreement and thus served as important fundamental theories in producing intelligence test.

Theoretical Framework of Intelligence

Many psychologists have employed several criteria and methodologies to study intelligence and one of the ways in determining the intelligence framework is factor analysis, a tool used in the analysis of data from the measurement of the ability of an individual in many aspects. Factor analysis uses statistics in classifying types of data that are related and is then applied in grouping similar individuals together. It involves
several methods with each one being different from the others in terms of details. Each method begins with the analysis of the correlation among variables. After analyzing the various factors, the interpretation of the meaning of each factor is followed depending on the knowledge about theories on psychology. The group of psychologists who used the technique of factor analysis as a tool to analyze data, had created a theoretical framework of intelligence, as presented below.

Spearman (17), who was the first researcher to use factor analysis to analyze data on psychology, proposed a Two-Factor Theory by conceptualizing that the intelligence ability of an individual consist of two elements, namely: General Factor or the g-factor which is the fundamental ability of an individual to use reasoning as mainly influenced by heredity thus causing individuals to have different levels of g-factor; and the other element, the Specific Factor or the s-factor, which is a specific ability in doing each activity and is greatly affected by education. When both specific factor and general factor are combined together, they become an indicator of success of the activities. Aside from these, Vernon (18) supported Spearman’s theory by using this theory to study and further develop it until the Hierarchical Theory was created, and which refers to intelligence as a brain activity that is divided into two types: Fluid Intelligence and Crystallized Intelligence. In doing different activities, the human brain has to utilize factors of verbal education and practical education to do all the activities. This also fitted into the theory of Cattell (19) which divided intelligence into two parts: fluid intelligence which is the ability of the brain due to genetics or anything except related to experience or learning; and, crystallized intelligence as an ability due to learning and experience about anything such as the ability to use languages, calculation, or solving a problem, which receive learning until these abilities are created.

Aside from the above mentioned theories, there are other several theories that showed intelligence having other important factors such as in 1927, Thorndike (20) presented the Multifactor Theory which states that intelligence was comprised of several types of specific abilities with each type of ability having an absolute dominant characteristic. In doing one intelligent activity, there is a need to use together several specific sub-abilities as a group. Thorndike divided these sub-abilities into 3 groups, namely: social intelligence which is the ability to adapt to be others; concrete
intelligence which is the ability in using one or more tools easily until they become intricate; and abstract intelligence which is the ability to learn abstractions such as semantics and mathematics. Besides Thorndike’s theory, another theory related to intelligence is the Guilford (21) theory. In 1959, he presented a theoretical framework of intelligence in a 3-dimensional model as an analysis of the relationship of the characteristics as contained in three factors: operation comprised of 5 types: cognition, memory, divergent thinking, convergent thinking and evaluation; content comprised of five types: visual, auditory, symbolic, semantic and behavioral; and products which come as a result of the process of an action comprised of 6 six types: units, classes, relations, systems, transformation and implications. When all these components of the three factors are joined together, the intelligence framework then contains a total of 150 factors with each factor having the ability to explain in 3-dimension the capability of each factor. One of the other important theories on intelligence is the theory of Thurstone (22) who, in 1983, proposed the group factor theory which stated that man’s ability is a primary mental ability of a person and has seven important components: on semantics or V-factor which is the ability in knowledge and understanding about the definition and relationship of words; on quantity or N-factor which is the speed and precision in easy mathematical calculation; on memory or M-factor which is the speed and precision in memorizing details of what was seen or learned; on word use or W-factor which the ability to use various words, synonyms or antonyms; on reasoning or R-factor which is the ability to summarize a criterion from the small to big aspect; on related dimension or S-factor which is the ability to create imaginations about graphics; and, on perception or P-factor which is the ability to recall things kept in memory. These 7 aspects are related to intelligence in terms of thinking and doing things.

These original theories were found to have both similarities and differences but each theory is able to explain intelligence in almost the same direction, indicating that intelligence is only a component that explains human adaptation towards various changes that occur in the surrounding environment. From this concept, it can be summarized that general perception or intelligence has two types of components, ability from learning and ability that is inherent within each individual, thus for
serving as conceptual basis for the creation of an intelligence test that contains both verbal and nonverbal parts.

**Development Theories on Intelligence**

The highly referred development theories on intelligence were formulated by 2 psychologists, namely: Piaget from Switzerland and Bruner from USA.

The theory of Piaget was widely accepted to explain the development of intelligence of a child which placed importance on cognitive theory that describes the development of a child’s intelligence as the ability to manage his environment and ability to perceive about the environment in consideration and understanding. Development takes place as a sequence and is similar to each child with growth at each stage being a composite of all behavior the child has learned in each stage. Aside from this, the child has a period of time for preparing to build his thinking and understanding and also a period time to manage to correct his past experiences before the ability to think or create a concept occurs. Piaget summarized the child’s intelligence in four stages (23) as:

The first stage occurs from birth to 2 years of age as sensory motor stage when both physical body and senses are used. The child at this stage develops his skills of the muscles and movement which is an important fundamental development of intelligence. The child attempts to accommodate himself to his environment and developing his ability to assimilate or accept the truth of his environment.

At 2-7 years old, the child starts to have thoughts and understanding. At this pre-operational stage, the child begins to understand different symbols and is able to use them efficiently. A two-year old child starts to understand the meaning of symbols that represent real things and also, begins to separate differences between those symbols. Meanwhile, a child of about 6-7 years old, develops his understanding and actual concepts of conserving the relationship of things.

When a child reaches 7-11 years of age, he begins to use his thinking with concrete reasoning. In this period called the concrete-operational stage, the child forms reasons for every thing that he has not understood before and also starts to understand the similarities of symbols and their contents and the changes of shapes to be either
with more weight or remains the same. The child at this stage is able to use his knowledge or past experiences to solve the problems in new events besides being able to transfer his learning. But those problems or events must be related to a thing or a concrete object. As for problems of abstract nature, the child is still not able to solve them. Ginsberg and Opper (24) further explained that this stage of Piaget’s theory as a period when a child starts to receive primary education thus there is development of his thinking and with reasoning based on logics in a concrete manner such as an object that he can grasp or hold. At this stage, the child begins to think decreasingly that he is the center of everything. He also thinks about the opinions of other people and starts to use popularly used language and responds with clarity. Besides these, he is able to sort out his many thoughts at the same with a method of solving the problems by using concrete reasoning with importance.

The age from 11-15 years is the period when a child thinks with abstract reasoning and is called the period of formal operation. The child starts to have the ability to think with abstractions, to create assumptions by using his imaginations or his own observations, and to have greater ability to compare and estimate. From the above, it can be seen that child development in that later stage of development is considered the pre-teen period when a child has wider intelligence, begins to get interested in many kinds of books, has initiative thoughts to try something, and has greater ability to think and solve problems. This can be seen from his ability to use reasoning, correct understanding of the spoken words and is able to limit in the words the abstractions. The child is increasingly able to see the relationship of things with a more improved memory and there is greater change in his thinking process according to his age (24).

Aside from these, another theory concerning the development of intelligence is Bruner’s theory. In 1961, Jerome Bruner reported about the development of perception and thinking in which some parts are related to Piaget’s theory. With similarity to Piaget, Bruner stated that the development of intelligence of an individual is based on the order and the emphasis on the importance of the environment that affects the development of the child’s intelligence. Bruner divided the development of human intelligence into 3 stages, namely: enactive stage which can be compared to sensory
motor period of Piaget when the child begins to learn more from doing things; iconic stage which is comparable to pre-operational period of Piaget when the child had greater realization of the truth and is able to think more from his perception together maybe with some imaginations that prevent him from having deeper thoughts, a similar stage with Piaget’s concrete operational period; and lastly, symbolic stage which is the highest development period of Bruner when the child has an understanding of things and is able to think comprehensively without any confusion. This last stage is similar to the concrete operational period of Piaget because it was the intention of Bruner who stated that these 3 stages of development sprout from the development within the physical body. Afterwards, the development of the senses and thoughts depend largely on the cultural environment of the individual.

From the theories on the development of intelligence in the different growth stages, it can be seen that the development of intelligence between 7 to 11 years of age is an important period because the child at this stage starts to enter the schooling period which provides greater development in his thinking process. The child begins to know how to use reasoning as a concrete means to solve problems besides learning from his past experiences to solve the problems caused by events at the present time. Also, the child learns how to plan ahead with highest imaginations and is able to think and remember either in playing or studying. Thus these different stages of development serve as essential and important aspects in measuring intelligence. In this study, because the measurement in TONI-3 and CPM intelligence tests depends largely on those abilities in testing intelligence, the researcher, therefore, opted to select children with ages ranging from 7 to 11 years because it can be seen that these ages are appropriate and proper to the objectives of the measurement using the tools used in the intelligence tests.

Properties of Psychology Tests

The suitability of intelligence tests used in measuring abilities were considered from the important characteristics of the psychology tests which consist of 3 elements: validity, reliability and practicality.
Validity

The validity of the intelligence test can be seen from a high quality intelligence test with validity being the most important characteristics that any kind of intelligence test should have because it can be seen that the type of intelligence test is able to measure the ability of the person actually giving the test or able to measure the validity based on his need to measure. Cherdsak Kowasin (25) stated that validity refers to the property or ability of the measuring tools to measure what needs to be measured, whether it is within the scope of the content of the operational definition or framework in similarity with the operational definition of Amnuay Lertchayandee (26) who stated that the validity of the tool that is used to measure performance is a quality assurance of the test scores or numbers from the measuring tools that have validity to be able to measure based on the objectives of the measurement. From that limitation, it can be summarized that validity means the property of the measuring tool that shows its ability to measure what it needs to measure based on the objectives of the test tools.

The testing of the validity involves several aspects. In 1996, the joint committees of the American Psychological Association (APA), American Education Research Association (AERA) and the National Council on Measurement in Education (NCME), classified validity into 3 types, as (5):

1. Content validity, which refers to the ability of the intelligence test to measure in fitness to the purpose of the intelligence test and divided into logical validity which uses the expert’s consideration and face validity considers that whether each intelligence test is valid based on its operational definition or not.

2. Criterion–related validity is a determination of validity that uses test scores to find the relationship with external criteria and which could be divided into two types: concurrent validity which determines the relationship with present criteria; and predictive validity which determines the relationship with future criteria, in order to use the test results to predict the results in the future.

3. Construct validity or as in theory, validity which refers to the ability of the test to be able to measure the behavioral characteristics based on the theory or framework as prescribed. This is also a type of validity that is suitable with tests that
have contents and behaviors that are more unclear than tests that measure assumptions which have contents and behaviors that are already clear.

There are several factors that affect validity such as the factor that is inherent to the individual being tested and who has no knowledge and understanding of the test, has temper and physical abnormalities and unpreparedness. Aside from these, some may be due to factors in the implementation of the test which may not be standardized and considered as unclear, no strict supervision of the test, timing is not precise, there might be mistakes in its implementation, and allocated time is either more or less than enough including factors due to the characteristics of the test which may be considered difficult, contains questions that are confusing or the test may be too short (27). These make the various standards of the validity of the test may be different, affecting the resulting scores to be unbelievable. Thus, in the implementation of the different tests, there is a need to determine the validity of the test. In this study, the validity of the TONI-3 intelligence test is determined particularly the content-related validity, by comparing with the CPM standard intelligence test, which is an intelligence test used to study and is diversely used in many areas. On the other hand, it is an intelligence test that is standardized and is widely accepted in terms of statistics of researchers in education both in and out of Thailand and also, there has already been wide identification of normal criteria and validity of this test in the sample population consisting of Thais.

Reliability of Intelligence Tests

Reliability is a characteristically important aspect of the research tool in psychology which aside from the need from already having validity in testing, must also have reliability because reliability shows the correctness and precision of the measurement process that makes the results or scores have a level of consistency without varying or changing due to change in time too. If there is no resulting effect towards the change of the variables that impact the data respondent during the different times of measurements (27), similar to the definition of Ratana Siriphanich (28) that reliability is the consistency or certainty in measuring the same object at different times. The determination of the reliability of the intelligence test is a
A mathematical process to serve as an index that indicates that an intelligence test has a sufficient value of consistency of its measurement. If measurement is done several times, its consistency accepted to be similar in each aspect. It could then be summarized that reliability is a characteristic of an intelligence test which has the same results as the original and with no change even if measured in different times with each measurement being independent of each other.

However, it was found that the measurement of the tools contained errors that always provided interferences thus there was a need to determine the reliability coefficient in order to absorb the errors of that tool. In this study of determining the reliability, different methods are applied such as the test-retest reliability, which is the use of the tools twice at different times on the same sample population and then using the results of the measurement to find the correlation value. The other method called the parallel form is also the use of two sets of tools which are comparatively similar to a common sample population. The method of determining the correlation value is similar to that of the test-retest coefficient. Another method used in determining reliability value is the measure of internal consistency which is a way to measure reliability from one set of data collection followed by the selection of the method of analysis between Kuder-Richardson which analyzes the scores from the sub-items and then dividing them into separate items, and Cronbach’s Alpha which has a base from the calculation of dispersion of each sub-item of the intelligence test (25, 29). Another concept is the split-half reliability method which tests the sample group which is used to measure the samples together and then analyzing the scores while dividing them into two similar groups from the test with similar content and ease-difficulty.

Even though the determination of the reliability uses several methods and each is different from the other, but everyone has a similar goal of determining the results of the measurement from the intelligence test at different time periods with the results do not undergo changes and each measurement is independent thus it can be said that reliability is an important property of testing tool.
Practicality of Intelligence Test

Besides the validity and reliability being considered important properties of the intelligence test, there is still another property that has similar importance. Practicality is another property of the test tool on management which means that the tool can be put to actual use with consideration on the economy in terms of time and cost in the implementation of data collection and convenience in the use of the tool, aside from interpretability which means that when other individuals who are not researchers use this tool, they are able to interpret the meaning of the results of the measurements as similar to that of the researchers (25).

From the 3 types of properties of the test as previously described, the researcher has seen them to have importance for the study of each testing tool because these are the properties that make the tests credible and precise based from the results of the test. Thus a good testing tool must be able to measure directly the ability of the individual based on the status and comparison with the value at present. That test tool must be considered to have real quality so in the use of a quality standard test, there is an initial need to study its properties. In this research, the criterion-related validity of TONI-3 was studied in comparison with CPM which is a standard intelligence test that is commonly used by researchers both in Thailand and abroad.

The Non-Language Intelligence Test

Definition of Language Free Test

In using the language free intelligence test, the concept that explains the implementation of intelligence test other than original non-verbal test, indicates that the content of the intelligence test together with the directives of the language free test does not use words of spoken language during the entire implementation of the test except the non-verbal language only because the originator of this test realized that the implementation of the verbal test is not able to measure the level of ability of the person being tested due to the limitation posed by language or cultural differences (30).
Sax (31) defined language-free test is the implementation of a test that includes the explanation and advice to the person being tested by using communication that is non-verbal or through gestures, which is confirmed by Anastasi (6), Aiken (32) and Hammill (33) who reported about non-language test as a test that contains non-verbal wherein both the person giving the test and the one receiving the test do not use any verbal language at all during the test.

**Test of Nonverbal Intelligence, Third Edition (TONI – 3)**

The TONI-3 intelligence test was developed by Brown, Sherbenou and Johnsen with the objective of measuring the person’s ability in terms of his senses, perception and use of reasoning. It was originally created in 1982 and was named the Test of Nonverbal Intelligence (TONI). At first, it was used to study particularly the sample population comprised of those who have writing problems, deaf and are handicapped only. Thus in TONI-2, the population is increased to include those who cannot speak the English language or have limitations in the use of this language, those who have abnormalities in reading and those with good intelligence, thus referred to as Test of Nonverbal Intelligence, Second Edition (TONI-2). Later it was improved once more in 1997 and was called the Test of Nonverbal Intelligence, Third Edition (TONI-3) (30), an intelligence test that is used at present and has been suggested to use in this study. The improvement on this test was on the reduction of the number of questions to 45 perform and on the increase in the age of individuals receiving the test from 5 years to 85 years and 11 months into 6 years to 89 years and 11 months. From these improvements, the norm for the new sample population was determined to attain more credibility and acceptability.

The TONI-3 is a language-free intelligence test which means that there is no use of verbal language during the entire implementation of the test thus being able to solve problems on the difference in language and culture of the individuals being tested. As such, this test can be used to individuals with language deficiency in terms of reading, writing, speaking and listening to the English language. On the other hand, this test is used for persons with excellent language skills but with no proficiency in
writing and speaking English. This test is used for individuals with ages from 6 years to 89 years and 11 months and consists of 2 papers called Form A and Form B. Questions in this test are presented in graphics form of different types. Each item has a picture that lacks some pieces and the person being tested must select one answer only from 4-6 choices. The correct answer is related to the graphics and the 45 questions per form are arranged in order of increasing difficulty.

The TONI-3 intelligence test was created in order to measure the ability to use reasoning in abstract thinking together with the human brain functions in solving problems. This is in conformity to the theory of Spearman which states man’s ability in terms of g-factor in which Spearman believes that g-factor is an important fundamental element of human sense in solving problems particularly in the perception through sight by identifying the relationship various things comprehensively and logically. The creator of TONI-3 used this concept together with the theory of Thurstone which emphasizes the importance of the basic components of human brain in 7 aspects, namely: semantics, mathematics, memory, vocabulary, reasoning, relative dimension and perception. It was found that the ability in these 7 aspects serve as basis from the g-factor component based on the theory of Spearman.

In the examination of the calculation of scores from TONI-3 intelligence test, a score of 1 is given for a correct answer. This is the raw score and is later calculated as a standard score. Later, the percentile ranking is allocated for each age to know the results of intellectual measurement from answering the test with this intellectual level from the test classified according to percentile ranking into 7 levels, as follow:
Table 1  Classification of standard score by the TONI-3 test

<table>
<thead>
<tr>
<th>Description</th>
<th>Percentile Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very-superior</td>
<td>More than 98</td>
</tr>
<tr>
<td>Superior</td>
<td>91-98</td>
</tr>
<tr>
<td>Above –Average</td>
<td>74-97</td>
</tr>
<tr>
<td>Average</td>
<td>25-73</td>
</tr>
<tr>
<td>Below-Average</td>
<td>9-24</td>
</tr>
<tr>
<td>Poor</td>
<td>2-8</td>
</tr>
<tr>
<td>Very-Poor</td>
<td>Less than 2</td>
</tr>
</tbody>
</table>

Norms of TONI – 3

In TONI-3 intelligence test, criteria for norms have been created using a sample population of 3,451 individuals from 28 states in the United States of America (30) by dividing the study into two groups: 6-18 years old school children and 19-89 years old adults. In this study the demographic characteristics of the respondents were investigated, which for adults included hometown, gender, citizenship, nature of residence (urban, rural), nationality, nature of disability, family income, and educational level. Meanwhile the children were divided into 4 growth stages while adults were divided into 6 growth periods with criteria normally studied to include 3 types, namely: standard score, percentile ranking and age equivalent. Based on the analysis of the collected data, it was found that average standard score was equivalent to 100 and standard deviation was 15. Afterwards, the standard scores were changed into percentile rankings and raw scores were changed into age equivalents.
Validity of TONI – 3

The validity of a test is a property that indicates that test measures precisely what it needs to measure based on the objectives or its validity may perhaps be the correlation value between the test and the criteria. When considering the 3 types of validity: content validity, criterion-related validity and construct validity (25), the validity of TONI-3 can be explained in details as follow:

1. The TONI–3 intelligence test has content validity because of 3 reasons: firstly, there was a choice in the form and questions of each sub-test as justified by the content of the intelligence test; secondly, in each question of the intelligence test was selected from the development stages in order to conform with the structure of the test; lastly, in each question, analysis was made to prevent prejudice in each question (30).

2. The TONI–3 has validity in terms of the criterion-related validity similarly with Hammill, Pearson and Wiederholt (33) who studied the relationship between the Comprehensive Test of Nonverbal Intelligence (CTONI) and TONI–3 and who found the criterion-related validity of the test because the results had high correlation (a coefficient value of .76 in Form A and .74 in Form B). This showed that these two tests were able to measure intelligence without using language. Aside from these, a study on the relationship between TONI-3 and a verbal test, Wechsler Intelligence Scale for Children- Third Edition (WISC-III), showed a high correlation (a coefficient value of Full Scale IQ in Form A at .63 which was similar to Form B), and could be compared with the correlation between TONI-3 and a verbal intelligence test called the Wechsler Adult Intelligence Scale-Revised (WAIS-R). It was found that there was also a similarly high correlation (a correlation coefficient of Full Scale IQ in Form A at .73 and in Form B at .71), which showed that TONI-3 test was able to measure intelligence without verbal language as with verbal test.

3. The construct validity referred to the ability of the test to be able to measure the characteristics and behavior based on the theories or constructs as indicated by the creating the TONI-3 test based on the fundamental theories of intelligence consisting of 6 aspects and which allowed the analysis of the validity based on the related constructs of the six aspects, as follow:
1. On Age Differentiation, which referred to the development of intelligence that started ever since childhood and developed continuously to the early adulthood (about 30 years of age) until the highest peak reaching the seniority age when it begins to decrease slowly in direct proportion to increasing age (34). The efficiency of TONI-3, therefore, must support the type of development of one’s intelligence all throughout man’s age. From the study conducted by the creator of TONI-3, it was based on the theories on the development of intelligence by studying the 20 periods of an entire age and it was found that the average raw score from the intelligence test became higher when age was increased from 6 years to 17 years. Afterwards, the average score almost remained at 30 in both Form A and Form B, until the age of 60 years and higher when the average raw score slightly decreased. From the study at the age 6 to 17 years and 11 months, it was found that average raw score was highly correlated with age as shown by the coefficient value of .63 in Form A and .60 in Form B. When considering the age, results indicated that TONI-3 intelligence test had construct validity according to the theories of the development of intelligence.

2. On Correlation with School Achievement, because TONI-3 test used the measurement of aptitude, the results of the test should have relationship with success in education. The creator of this intelligence test expected that the results of the TONI-3 test should be positively correlated at a low to moderate level with success in education by determining the correlation between the TONI-3 test and the measurement of 5 components of success in education of 20 students. Results showed a correlation value from .55 to .76 (30) and a median value of .71 which led to the summary that there existed a correlation between TONI-3 test with measurement of success in education. By considering this, TONI-3 test had construct validity.

3. On Group Differentiation, TONI-3 test was a measurement of intelligence thus the results were different between the group with high and average levels of IQ and the group with lower and higher than average levels of IQ. Based on the results using 13 groups with differences such as 8 groups of various handicapped characteristics of both male and female children, 1 group with special talents, 2 groups using English language and 2 groups of Afro-American respondents, it was found that sub-groups of the individuals with abnormalities had lower standard deviation from
the average value and in the group of individuals with special talents, results showed the test scores were higher than the average score. As for groups with differences in language, citizenship and skin color, results indicated no difference with average standard deviation at about 90-110 which showed that TONI-3 test was able to control the differences in culture among individuals in various groups and results of the correlation with the theories on the differences among individuals in the group.

4. On Multiple Correlation and Regression, as stated previously that TONI-3 test is a form of measurement for intelligence thus the results of the measurement can be used to predict the efficiency of other intelligence tests by being able to study from multiple correlation and regression, however, there is no calculation of multiple correlation and regression in TONI-3 test where there is only a specific researcher for TONI-2 test in comparison with WISC-R test which showed that TONI-2 is able to predict well the results of the WISC-R Full Scale IQ test, therefore, it cannot be summarized that TONI-3 test can also predict the efficiency of other intelligence tests.

5. On Factor Structure, since TONI-3 is a form of intelligence test that was created to measure the use of abstract reasoning in solving problems, which is based from theories on the elements of intelligence by Spearman, thus each question of TONI-3 should be arranged by itself based on the larger components. When the factor analysis techniques are used which arise from the analysis of the factors of TONI-3 test in Form A and Form B for each question, results showed that the initial principal components of Form A and Form B tests had a variance of 67% and 68%, respectively, which are considered quite high and indicates that TONI-3 test for each question also measured the characteristics of the g-factor component and applying the method of Promax Rotated Factors, 3 components occurred with variance of 97% in Form A and 99% for Form B thus supporting the hypothesis of the test creator.

6. On Item Validity, because each question of this test similarly has measurement of the characteristics, thus it was found that each question should have a linear relationship. Results of the study with different questions showed a median discrimination coefficient value of .49 in Form A and .50 in Form B because these two types of tests were designed to have linear relationship.
Reliability of TONI – 3

From the study using the TONI-3 intelligence test for individuals with ages ranging from 6-89 years, it was found that the results of the test had an average reliability which was determined by the transformation of that reliability value to a standard value or Z-transformation following the method of Guilford and Fruchter by studying the content sampling which gave the reliability value of 0.93 in Form A and 0.92 in Form B. On time sampling, reliability value was 0.91 in Form A and 0.92 in Form B. And, on the reliability of scorer differences, reliability was 0.99 in Form A and Form B. Based on the reliability value of the 3 forms, reliability value for this test became 0.96.

With the characteristics and properties of the TONI-3 test as previously explained, the researcher considered the use of TONI-3 test as a tool in the classification of intelligence in the population group composed of Thais which consisted of sick persons or those who may have problems in communication in the future in Thailand. However, since there has been no study yet on the validity of the TONI-3 test for group of Thai people in which validity is considered as an important property of the intelligence test as described earlier. Thus the researcher was interested to study the criterion-related validity of the test in comparison with CPM test in order to determine whether TONI-3 test was suitable as a tool to select or sort out intelligence of the Thais or not.

Colored Progressive Matrices (CPM)

J.C. Raven, a British psychologist created the Raven’s Progressive Matrices in 1938 as a non-verbal intelligence test to measure the ability of an individual by determining the relationship between geometrical shapes. The characteristics of this intelligence test are presented in a matrix as in various designs. As the matrix of each question of the test is designed to be missing, the one being tested had to choose one item from 6-8 choices. The correct choice would then be able to fill in the missing item to complete it. The items are arranged from easy to difficult (4)
This intelligence test was created to measure the ability of an individual in the use of reasoning as an important component of a general intelligence based on the theory of Spearman called the g-factor by solving that arithmetic problem. The person taking the test must rely on his perception in relation to his various senses particularly his perception to focus together with the determination of the relationship as comprehensively and with reason, in which Spearman has stated that the ability of the component in g-factor has two aspects:

1. Educative, which refers to the ability to provide meaning to an unclear object, prediction, and understanding of the complicated construct of the problem and decision making on various issues.

2. Reproductive, which refers to the behavioral process of thinking as a result of language, knowledge of various issues and knowledge from culture.

Raven used this concept together with group theory of Gestalt in the determination of ability in solving problem from CPM test that is composed of analysis from perception through sight, relational dimension, comparison of the similarities of the changing characteristics, understanding of the meaning of differences, and creating the identity of that object, in a systematic manner.

From the study, it was found that the basis that the person is taking the test by selecting graphics to complete the sentence, was different according to each person. But when the study concerned with the over-all view, the summary can be made on the basis which the person taking the test used altogether into 6 bases, namely: differentiation, completion, use of reasoning, verification, advanced thinking and backward thinking.

These 6 bases refer to the intellectual ability of an individual as envisioned by Raven himself who then created all 3 progressive matrix tests, as follow:

The Standard Progressive Matrices (SPM) is used to test individuals from 12 years of age and for every educational level and citizenship.
The Colored Progressive Matrices (CPM) is used to test individuals with ages ranging from 5 to 11 years and for those who do not use the English language and with problems on the ability to speak due to brain disease or with listening problem or deafness.

The Advanced Progressive Matrices (APM) is a test used for individuals with ages from 12 years and generally, to individuals with intelligence level that is higher than the ordinary individuals.

The Colored Progressive Matrices (CPM), which was used as a testing tool for this research, is a colored set of test to children of 5-11 years of age including the handicapped and the aged. The test uses printed colors in order to make the test more interesting without using many words. The test has 3 sets, namely: A, Ab and B. Each set is different from the other in terms of the difficulty of the questions. The form of illustrations in set A is less difficult than set Ab which is also less difficult than set B. Each set has 12 questions and each question has 6 choices. If the child is able to complete set A, this means that he has the ability to continue to the next test. If the child is also able to finish set Ab, it means that he has the ability to see the relationships of the graphics that have been separated from the big picture. As for set B, these are the problems that show the ability of the child to think with abstractions as cause and effect.

In the examination of scores, 1 is given to a correct answer and all scores are summed up together and compared to the percentile ranking for each age range. This would provide knowledge on the result of the test which is divided into 5 levels, as indicated in Table 2.
Table 2 Classification of the CPM test

<table>
<thead>
<tr>
<th>Description</th>
<th>Percentile Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very-superior</td>
<td>More than 95</td>
</tr>
<tr>
<td>Superior</td>
<td>75-94</td>
</tr>
<tr>
<td>Average</td>
<td>25-74</td>
</tr>
<tr>
<td>Low-average</td>
<td>5-25</td>
</tr>
<tr>
<td>Deficient</td>
<td>Less than 5</td>
</tr>
</tbody>
</table>

Validity of CPM

Wilkes and Weigel (35) studied that CPM test has a correlation coefficient with WISC-R test at 0.67 while Suchira Phattarayutawant, et al (7) reported that CPM is a test that measures the cognitive ability of an individual and on perception, which are contained in the g-factor of Spearman. Aside from this, Wutisak Nimmalangkul (36) stated that in the examination of the measurement of intelligence, the correlation coefficient with DAP test was at moderate level (r = .448).

Reliability of CPM

Valencia (37) studied the reliability of the Colored Progressive Matrices test of Raven using a group of 96 children of Anglo and Mexican American heritage who were currently studying the 3rd grade and who had lower economic and social status. Results of his study indicated that the assessment of the reliability of the CPM test administered to children with mixed citizenship, produced high reliability value.

Suchira Pattarayutawant, et al (6), who established the normal criteria of the Colored Progressive Matrices test using the sample group of Thai individuals and calculated the reliability value using the Kuder – Richardson (20), found that r was equivalent to .91.
Aside from this, another study was also conducted on the quality of the CPM test among children with ages ranging from 6 to 11 years by Siri Udmaphol (7) in the northern region and who found a high reliability of .92. Likewise, Benjawan Ransaewa (8), who conducted the study in the southern region, found a high reliability of .91 while Sudarat Sirisakphanich (9) reported a similar high reliability of .91 in the central region. Janthanee Mungkhetklang (10) also made the same study in Bangkok Metropolis and found a high reliability of .90 while in the northeast region, Wararat Inthabthim (11) attained a similar high reliability of .92 while Vichaya Mopheechat (12) obtained a high reliability of .92 in the southeast region. Lastly, Sureerat Palahkat (13) also attained the same high score of reliability at .92. All these values together with the national norm, had a total reliability of .92 (38).

From the above data, it can be seen that in the past CPM test was studied and used diversely in several sites aside from being a form of intelligence test that was standardized and widely accepted in terms of statistical researches both in local and foreign settings. The results of those initial studies showed that CPM test is a standardized test that can be used to sort and classify intelligence of the group of Thai individuals thus the research chose to use the CPM test to serve as a standard criterion in the study of the validity of TONI-3 intelligence test which is a test that has been less used in Thailand.

Conceptual Framework

The TONI-3 test was created in order to measure the ability in solving problems by the use of abstract thinking together with the function of the brain in perception so as to solve problems, which is based from the theory of Spearman in terms of g-factor and from the theory of Thurstone in relation to group components as indicated similarly by the theory of Spearman. Aside from this, it also included the theory of Piaget who reported about the maturity in those concepts and with TONI-3 being able to sort or classify intelligence in terms of these various aspects. Based on these, the researcher studied on the measurement properties of TONI-3 test by using CPM test as a standard criterion in the study of the quality of TONI-3 test since the CPM test has properties that are suitable to be used as criterion in the comparative
study especially because in terms of validity, CPM has already been studied to have validity in comparison with other tests using sample groups of Thai individuals which were related and were studied to determine the norm of the group of Thai population in each age group. Besides, there were still many other researches involving CPM test which showed that when CPM test was used to measure intelligence of the sample group of Thai individuals, reliability value was high. This led to the researcher initially formulating the hypothesis that these two intelligence tests were correlated at a moderate level because they had some parts that were different and other parts to be similar for use in measuring intelligence.

In this study, the researcher was interested in studying the validity of the TONI-3 intelligence test in terms of age differentiation in order to determine the differences in each age level and multiple correlation to determine the relationship with a standardized intelligence test such as CPM in which the researcher has expected that the TONI-3 test would serve as one other tool that can sort or classify the level of intelligence of Thai children because of its characteristics and properties as previously explained.
Figure 1 Conceptual framework of the study
CHAPTER III
MATERIALS AND METHODOLOGY

As a descriptive research, this study focused on the validity of one type of an intelligence test called the Test of Nonverbal Intelligence, Third Edition (TONI–3), using a population group of school children with ages ranging from 7 to 11 years who were currently studying in 2007 in a grade school under the Commission on Basic Education, Ministry of Education in Zone 3-Bangkok. The study was conducted in the following steps and methods:

1. Population and sample
2. Materials of the Study
3. Data collecting
4. Data analysis

Population and Sample

Population

In this particular research, the study population consisted of school children, both male and female, whose ages ranged from 7 to 11 years in Zone 3 – Bangkok in 2007, and were selected from a large school because of the large number of school children who could participate in this research to make it more convenient to select a sample group thus this group of school children to become a much better sample group from the study population than a medium or smaller size school.
Sample group

The sample group used in this study came from a simple random sampling which was done using the following step:

Step 1  Determination of the distribution in the selection of sample group and power of statistics

Let $e = 0.10$ or 10% of standard deviation (SD) : $Z = 1.96$

Estimate the number of school children with ages between 7 and 11 years from the 2006 operational database of the Office of the Commission on Basic Education, Ministry of Education, which showed that the number of school children in Zone 3-Bangkok who were 7-11 years of age was 7,571 children (39).

Formula 

$$n = \frac{NZ^2\sigma^2}{(N-1) e^2 + Z^2\sigma^2}$$

$$n = \frac{(7571)(1.96)^2}{(7571-1)(.10)^2+(1.96)^2}$$

$$n = 365.67$$

thus, based on the computation, the sample group should not be less than 366 children.

Step 2  Simple random sampling from 54 schools thus selecting Wat Amarinthararam School from Bangkok Noi, a school with large population for convenience in data collection.

Step 3  Sampling of the number of school children using statistical methods based on the number of school children from Wat Amarinthararam School whose ages ranged from 7 to 11 years studying in grade 1 to 6. From a total of 1,253 children, the sample group was selected to comprise 380 children from 2 rooms per grade level whose ratio of the number of school children was almost similar to the desired number.
Table 3 Number of sample group as distributed in gender and age

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>172</td>
<td>45.3</td>
</tr>
<tr>
<td>Female</td>
<td>208</td>
<td>54.7</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7:00-7:11</td>
<td>52</td>
<td>13.7</td>
</tr>
<tr>
<td>8:00-8:11</td>
<td>52</td>
<td>13.7</td>
</tr>
<tr>
<td>9:00-9:11</td>
<td>85</td>
<td>22.1</td>
</tr>
<tr>
<td>10:00-10:11</td>
<td>59</td>
<td>15.5</td>
</tr>
<tr>
<td>11:00-11:11</td>
<td>132</td>
<td>35.0</td>
</tr>
</tbody>
</table>

Research Instruments

1. Test of Nonverbal Intelligence, Third Edition (TONI – 3), an intelligence test that has independence in language and need for the response in the part of its movement which is considered less besides its fairness in culture as created by L. Brown, Sherbenou and Johnsen in 1997. This test is used on individuals whose ages range from 6 years to 89 years and 11 months and those who have abnormalities in speaking and writing, have auditory problems or controlling the listening ability, those who are not able to speak, read or write and those with no control in their movements. This test consists of 2 forms with parallel difficulties. The researcher used Form A in the initial test and Form B in the second test and each form was comprised of 45 questions and must be answered in about 15-20 minutes. The person being tested must read the questions from the question and answer sheet and afterwards, must record his answer on the answer sheet, with 1 point given for each correct answer, and must then stop when he recorded 3 consecutive wrong answers or 3 wrong answers out of 5 consecutive questions.

2. The Colored Progressive Matrices (CPM), an intelligence test which has culture fairness and non-verbal, was created by J. C. Raven, a British scientist, and
which was revised in 1956. This test is used for children whose ages range from 5 to 11 years and handicapped individuals. Three sets comprised this test, A, Ab and B with each set having 12 questions for a total of 36 questions. This test must be completed within 30-40 minutes with the individual being tested receiving 1 point for a correct answer but must answer the 36 questions.

**Data collecting**

1. Preparation for data collection

   The researcher practiced to master the implementation of the test and calculation of test scores from TONI-3 and CPM as supervised by the instructor and from the manual of each test.

   The researcher contacted the school through its academic affairs section for sending a self-introductory letter through the Graduate School of the university and indicated the objectives of the request to collect data for this research.

2. Data collection

   The research met with the school children and introduced herself and explained the research. Afterwards, the researcher distributed a document for permission to be a part of the sample group which the children should present for their parents’ consideration whether to participate in the research or not.

   After receiving the signed acceptance letter from the parents, the research started to collect data according to the study time table by coordinating with the school’s academic section on the schedule of data collection. The research gathered the data by herself but had an assistant who was made to understand about the method and steps in data collection, which was divided into two parts: group test and individual test, in order to determine the suitability of using this test.

   **Step 1  Implementation of the group test during September in the 1st school semester (2006)**

   The researcher created good relations before explaining the objectives of the test and the method of conducting the test to the sample group to enable them to have clear understanding. Afterwards, the conduct of the test was started with TONI-3 and
after each individual has completed the test, the implementation of the CPM test followed.

Step 2  Implementation of the individual test during February in the 2\(^{nd}\) school semester (2006)

As in the group test, the researcher created good relations before explaining the objectives of the test and the method of conducting the test to the sample group to enable them to have clear understanding. Afterwards, the conduct of the test was started with TONI-3 and after each individual has completed the test, the implementation of the CPM test followed.

**Steps in data collecting**

Stage 1  (Implementation of the group test)

The first step started with the conduct of the TONI-3 test which was given to the sample group that was divided into groups of 5-10 individuals. Afterwards, the individuals to be tested looked at the questions from the test proctor who then administered the test by each item whose answer was recorded in the prepared answer sheet until all questions had been answered. The test must be finished within 35-40 minutes. After finishing this stage, the group took a break with the teacher afterwards the 2\(^{nd}\) test was then administered.

The second step was implemented using the CPM test on the same sample group with the test administered similarly with the first test but only within 25-30 minutes.

Stage 2  (Implementation of individual testing)

The first step involved the conduct of the TONI-3 test in which the individual being tested looked at the questions from the test booklet and the test was administered per item with the person giving the test being the one who recorded the answers on the prepared answer sheet. This was done for each item until stopping when the ceiling age of the individual being tested was reached. The test was implemented within 15-20 minutes. After the test was conducted for all the sample group then the individuals
being tested had a rest together with one instructor and afterwards, the second test was continued.

The second step was conducted using the CPM test for the same group of individuals being tested. The test was administered according to the manual of CPM intelligence test and must be completed within 25-30 minutes.

**Data Analysis**

After all data had been collected, they were then analyzed using the Statistical Package for the Social Sciences (SPSS), which was divided into two periods:

**Period 1** (Implementation of the group test)

1. Calculation of the average and standard deviation of the test scores from TONI-3 and CPM tests according to gender and age.
2. Use of Pearson’s Product Moment Correlation method in determining the correlation coefficient which was used to study the correlation of intelligence as measured from TONI-3 and CPM tests.
3. Computation of the difference of average scores from TONI-3 and CPM tests according to gender and age by using the t-test (independent sample t-test).
4. Calculation of the difference of average scores from TONI-3 and CPM tests, as classified according to age using the One-way ANOVA through the method of Scheffe.

**Period 2** (Implementation of the individual test)

1. Calculation of the average of the scores from the TONI-3 and CPM tests using the individual testing method.
2. Computation of the difference of scores in TONI-3 and CPM which were administered both as group and individual test using the t-test (independent sample t-test)
CHAPTER IV

RESULTS AND DISCUSSION

Results

The objective of this research was to study the validity of the TONI-3 intelligence test in comparison with CPM with the group of school children of ages ranging from 7-11 years from a school under the Commission on Basic Education, Bangkok Zone of the Ministry of Education. The intelligence tests were conducted in two groups: group test with 380 primary grades aged 7-11 years and divided into 172 male and 208 female children; and individual test with 60 children from 9 years to 9 years and 11 months. Results of the study are presented in tables with discussions which are divided into two parts:

Part 1    Validity of TONI-3

Part 2    Analysis of the variance of average test scores from TONI-3 based on gender and age
Part 1 Validity of TONI-3 intelligence test

1.1 Study on the relationship of TONI-3 test scores with CPM test scores

This study was focused on determining the criterion-related validity by finding out the relationship between the test scores from TONI-3 and CPM and also the construct validity in terms of age differentiation through the analysis of variance using the Pearson product moment correlation coefficient and analysis of fitness in the classification of intelligence level. Results of the study showed that test scores of TONI-3 with CPM were moderately correlated in linear direction as indicated by the Pearson product moment correlation coefficient value of .466 with a statistical significance of .001.

1.2 Fitness of classification of intelligence of TONI-3 test with CPM (group test)

Table 4 Classification of intelligence (IQ) as calculated from composite scores of TONI-3 and CPM group tests

<table>
<thead>
<tr>
<th>IQ Scores*</th>
<th>TONI – 3 (%)</th>
<th>CPM(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>69 and below</td>
<td>Extremely Low</td>
<td>8.9</td>
</tr>
<tr>
<td>70-79</td>
<td>Borderline</td>
<td>5.5</td>
</tr>
<tr>
<td>80-89</td>
<td>Low Average</td>
<td>13.4</td>
</tr>
<tr>
<td>90-109</td>
<td>Average</td>
<td>42.1</td>
</tr>
<tr>
<td>110-119</td>
<td>High Average</td>
<td>10.3</td>
</tr>
<tr>
<td>120-129</td>
<td>Superior</td>
<td>8.9</td>
</tr>
<tr>
<td>130 and above</td>
<td>Very Superior</td>
<td>10.8</td>
</tr>
</tbody>
</table>

* distribution of IQ scores according to Educational use (30)
From Table 4, it can be seen that most of the school children attained IQ scores from the two tests according to Educational use at an average level.

However, even if majority of the scores of the school children were found in almost the same level, about 14.4 percent of the sample group obtained scores that were much lower than average in the TONI-3 test while in CPM, only 2.7 percent of the sample group had very low scores as compared to the normal score. It was also found that when testing was done on TONI-3 test, about 19.7 percent of the sample groups had score levels that were higher than the normal level. At the same time, when CPM was used, only 2.4 percent of the sample group had higher scores than the normal level.

**Table 5** Result of the analysis of fitness in the classification of intellect of the TONI-3 in comparison with CPM in a group test

<table>
<thead>
<tr>
<th>IQ Level</th>
<th>Fitness</th>
<th>Percentage</th>
<th>Non-Fitness</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>69 and below</td>
<td>3</td>
<td>8.8</td>
<td>31</td>
<td>91.2</td>
</tr>
<tr>
<td>70-79</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>100.0</td>
</tr>
<tr>
<td>80-89</td>
<td>5</td>
<td>9.8</td>
<td>46</td>
<td>90.2</td>
</tr>
<tr>
<td>90-109</td>
<td>106</td>
<td>66.3</td>
<td>54</td>
<td>33.8</td>
</tr>
<tr>
<td>110-119</td>
<td>19</td>
<td>48.7</td>
<td>20</td>
<td>51.3</td>
</tr>
<tr>
<td>120-129</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>100.0</td>
</tr>
<tr>
<td>130 and above</td>
<td>0</td>
<td>0</td>
<td>41</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>133</td>
<td>35</td>
<td>247</td>
<td>65.0</td>
</tr>
</tbody>
</table>
From Table 5 which showed the group classification of the IQ level, it was found that TONI-3 test was able to classify the IQ level into 7 groups although detailed examination of the percentage of fitness in the classification of the intellectual level, showed that the group with the highest fitness was the 90-109 (average) level at 66.3 percent. This was followed by the 110-119 (high average) level which was 48.7 percent. As for IQ levels of 70-79, 120-129 and 130 and above groups, results showed that no fitness was found which indicated that both tests were different in classifying the sample group because they had only 35 percent fitness. This was due to the high level of non-fitness. The researcher has thus considered this as a factor that affected the scores that were analyzed. It was observed from data collected that the characteristics of the group test of TONI-3 made it unsuitable for implementation. As such, there was a need for the implementation of data collection using individual testing to be repeated again by selecting a sample group of children with ages from 9 years to 9 years and 11 months, as shown by details presented in Tables 6 and 7.

1.3 Analysis of the differences between grouping test and individual test using Toni-3 and CPM test

Table 6 Classification of IQ level as calculated from composite scores of TONI-3 and CPM tests in both individual and group tests using 60 school children with ages ranging from 9 years to 9 years and 11 months
Table 6 shows the classification of the intellectual level of TONI-3 and CPM tests in both group and individual tests. It was found that when TONI-3 was administered through individual test, the classification was different from the group testing with the group testing being able to classified the sample group into 7 levels.

<table>
<thead>
<tr>
<th>IQ Scores *</th>
<th>Grouping</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classification</td>
<td>Percentage</td>
</tr>
<tr>
<td>69 and below</td>
<td>Extremely Low</td>
<td>4</td>
</tr>
<tr>
<td>70-79</td>
<td>Borderline</td>
<td>4</td>
</tr>
<tr>
<td>80-89</td>
<td>Low Average</td>
<td>3</td>
</tr>
<tr>
<td>90-109</td>
<td>Average</td>
<td>31</td>
</tr>
<tr>
<td>110-119</td>
<td>High Average</td>
<td>4</td>
</tr>
<tr>
<td>120-129</td>
<td>Superior</td>
<td>5</td>
</tr>
<tr>
<td>130 and above</td>
<td>Very Superior</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IQ Scores *</th>
<th>CPM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extremely Low</td>
</tr>
<tr>
<td>70-79</td>
<td>Borderline</td>
</tr>
<tr>
<td>80-89</td>
<td>Low Average</td>
</tr>
<tr>
<td>90-109</td>
<td>Average</td>
</tr>
<tr>
<td>110-119</td>
<td>High Average</td>
</tr>
<tr>
<td>120-129</td>
<td>Superior</td>
</tr>
<tr>
<td>130 and above</td>
<td>Very Superior</td>
</tr>
</tbody>
</table>
But when individual testing was done, it was only able to classify the sample group into 5 levels with no child having an intellectual level that was lower than Low Average. When the CPM test was conducted, results showed that individual testing similarly led to the classification of the sample group into 5 levels but 2 intellectual levels were different. The TONI-3 test was able to classify a high level of Very Superior and classified into the lowest level, Low Average. Meanwhile, the CPM test was able to classify the sample group into the highest level of Superior and the lowest level at Borderline. When considering the results of the test for individual testing of TONI-3, it was found that 53.4 percent of the sample group had a high IQ level of High Average and above. This was almost similar to CPM test which showed that most of the sample group (88.3%) had an IQ level of High Average and above.

<table>
<thead>
<tr>
<th>Method of Testing</th>
<th>IQ Value</th>
<th>Group Test</th>
<th>Individual Test</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M    SD.    M   SD.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TONI-3</td>
<td>103.88</td>
<td>20.45</td>
<td>109.95</td>
<td>13.91</td>
<td>2.326</td>
</tr>
<tr>
<td>CPM</td>
<td>106.43</td>
<td>9.04</td>
<td>105.13</td>
<td>13.22</td>
<td>-0.711</td>
</tr>
</tbody>
</table>

Table 7 showed the difference in the IQ values resulting from the two types of intelligence tests as administered by group and individual testing. It was found that when TONI-3 was administered in group testing, the sample group had IQ values that were significantly different at a statistical level of .05. When administered as an individual test, the IQ value of the sample group was higher than when using the group test. This showed that TONI-3 test was more suitable to be administered as an individual test than as a group test as indicated by the test manual. When CPM test
was administered in group or individual testing, it was found that the IQ values were not different which showed that the CPM test could be used as a group or individual test. When comparison was made on the results of the classification of the IQ levels, it was found that TONI-3 and CPM tests showed the results in the form of average score of the sample group had a similar IQ level of Average.

**Part 2 Analysis of the differences of the average score level from TONI-3 as classified according to gender and age**

**Table 8** Results of the analysis of the score difference between TONI-3 and CPM based on Composite Scores as classified according to gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>Total Score</th>
<th>Male</th>
<th>Female</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD.</td>
<td>M  SD.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TONI-3</td>
<td>19.06 9.78</td>
<td>19.86 9.29</td>
<td>-0.812</td>
<td>.166</td>
<td></td>
</tr>
<tr>
<td>CPM</td>
<td>30.91 4.78</td>
<td>30.13 5.54</td>
<td>1.47</td>
<td>.056</td>
<td></td>
</tr>
</tbody>
</table>

Table 8 showed the comparison of scores resulting from TONI-3 and CPM tests between male and female students. Results showed that based on the total score from TONI-3, the probability value (p-value) was .166, which indicated that even though the sample group was composed of different genders, the scores from TONI-3 was not significantly different at a statistical level of .05. As for the scores from CPM test, the probability value (p-value) was equivalent to .056, which showed that the school children who had different genders had scores from CPM test to be similarly not significantly different at a statistical value of .05.
**Table 9** Results of the analysis of the difference in scores from TONI-3 as classified according to age

<table>
<thead>
<tr>
<th>Group / Age</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/7:00-7:11</td>
<td>52</td>
<td>11.83</td>
<td>7.62</td>
<td>14.249***</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>2/8:00-8:11</td>
<td>52</td>
<td>17.96</td>
<td>7.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/9:00-9:11</td>
<td>85</td>
<td>19.21</td>
<td>8.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/10:00-10:11</td>
<td>59</td>
<td>22.53</td>
<td>8.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/11:00-11:11</td>
<td>132</td>
<td>21.94</td>
<td>10.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9 showed that when the variance of each age group of the sample group was tested, it was found that the scores of the sample group tended to have a probability value (p-value) of less than .001. This meant that the sample group which had different age levels had test scores from TONI-3 to be significantly different at a statistical level of .001 and when considering the average score values as indicated in the table, it was found that the average score of the sample group was increased when age level was increased. This could be seen clearly in groups 1, 2 and 3. When ages of the sample group were paired, results can be shown in Table 10.
Table 10: Results of the analysis of the difference in scores in age pairs at different age levels from TONI-3 test

<table>
<thead>
<tr>
<th>Age</th>
<th>7:00-7:11</th>
<th>8:00-8:11</th>
<th>9:00-9:11</th>
<th>10:00-10:11</th>
<th>11:00-11:11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 / 7:00-7:11</td>
<td>-6.1346*</td>
<td>-7.3874*</td>
<td>-10.6985*</td>
<td>-10.1129*</td>
<td></td>
</tr>
<tr>
<td>2 / 8:00-8:11</td>
<td></td>
<td>-1.2527</td>
<td>-4.5639</td>
<td>-3.9783</td>
<td></td>
</tr>
<tr>
<td>3 / 9:00-9:11</td>
<td></td>
<td></td>
<td>-3.3111</td>
<td>-2.7256</td>
<td></td>
</tr>
<tr>
<td>4 / 10:00-10:11</td>
<td></td>
<td></td>
<td></td>
<td>.5856</td>
<td></td>
</tr>
</tbody>
</table>

*p<.05

Table 10 showed the score levels of the sample group from 7 years to 7 years and 11 months. The average scores were less for every age level and were significantly different at a statistical level of .05 and when considering the comparison of age pairings based on the method of Scheffe, it was found that particularly for group 1, the average scores were significantly different for each group at a statistical level of .05 with group 1 having average score that was lower than each other group. When comparison was made on the different groups, results showed that for different ages, the average scores showed no significant difference at a statistical level of .05.
Discussion of the results of the study

The results of the study showed the value of validity of the TONI-3 test in comparison with the CPM test, with detailed discussion as follow:

Based on the database of the sample group being studied, it was found that the number and the percentage of the sample group in terms of gender and age, showed that female children were slightly more than the male children. When considering the age level, most of the children (35%) belonged to the group of 11 years to 11 years and 11 months thus making this study group not a good representative of this study in terms of age but may serve as a good representative in term of gender.

Study of the correlation of scores from TONI-3 test with CPM test in a group test

In the analysis of the value of Pearson product moment correlation coefficient, it was found that the scores from TONI-3 test were correlated with the scores from CPM at a significant difference with statistical level of .001. It was observed that the 2 types of tests had linear relationship in a moderate level (r=.466).

Further results of the study showed that both TONI-3 and CPM tests were created by a concept from the same theory and when considering the contents of each test, it was found that TONI-3 consisted of the use of Free language, Maturity, Visual Perception and Abstract Thinking in responding to the test questions. On the other hand, CPM test made use of Verbal Introduction, Maturity, Visual Perception and Logical Reasoning in responding to the test.

It can be seen that both intelligence tests had some similar characteristics together such as Maturity and Visual Perception. These two tests need to use the ability of an individual in these 2 characteristics in answering the tests, which was indicated by Kirati Bannakulroj (40) who studied that children with ages from 6-7 years and above started to have developed their ability to visually perceive things and then followed by the development of their other perceptions.
Aside from these, these two tests have characteristically similar contents with both using 2 theories, namely: intelligence theory of Spearman (17) on the general factors (g-factor) which constitute the general intelligence ability of an individual that consist of Educative, the method of understanding the elements in the past and Reproductive, comprised of memory, semantics and ability to learn; and the intelligence theory of Thurstone (22) which is related to the ability to use reasoning.

**Fitness to classification of intelligence level of TONI-3 and CPM tests**

Aside from determining the correlation of both tests, results of the study on the fitness of classifying the level of intelligence or IQ as shown by Tables 5 and 6, indicated that TONI-3 test was able to classify the scores into 7 levels which had more fitness at the Average level of intelligence, while CPM which is a test used as a criterion for this study, was able to classify the scores into 6 groups, as compared to TONI-3 which was more able to classify the Very Superior level, an increase in the contents of Superior level of CPM.

As TONI-3 was more able to classify the level of scores of the individuals being tested than CPM, it might indicate that the individuals being tested by two kinds of test could have their IQ values classified as different which might have affected the level of intelligence of the individual. From this, it can be observed that TONI-3 test was more able to assort the level of intelligence at the Extreme than CPM and this might came as a result of the implementation of the test as in group thus causing the results to be more dispersed. Thus the researcher had to repeat the test again by modifying the test as an individual test in order to observe the differences in the results that occurred.

When the ability to classify intelligence was considered, however, it was found that TONI-3 had a property of classifying that was different from CPM. Although the results of the 2 tests showed that most of the children had IQ at an Average level, but observing the percentage of the sample group who had IQ at the Average level showed a big difference in the values with the TONI-3 test having 42.1 percent. Meanwhile, in
CPM test, a higher percentage of 58.2 was found to have IQ at the Average level, showing a difference of 16.1 percent. When observing the IQ distribution of the sample group resulting from both tests, there was a difference in distribution with TONI-3 having individuals with IQ at Low Average as second and IQ at Very Superior level as third in the classification. In comparison with the results from the CPM test, the sample group had IQ at High Average as second in classification and IQ at Low Average as third. When considering the distribution of IQ level of the sample group based on the results from CPM test, it showed a curved characteristic and when comparing the fitness in the classification of the IQ level from the 2 tests, it was found that fitness was 35 percent with IQ at the Average level having the highest fitness at 66.3 percent followed by High Average (48.7%). IQ level at Borderline and Superior showed an absence of fitness in the classification which showed that both tests are still different in the classification of the intelligence level of children even though both are supported by similar fundamental theories.

Analysis of the difference of average scores from TONI-3 and CPM based on gender and age

In this study, a comparison was made on the average score levels of the sample group using both tests with classification based on gender and age. Since the dispersion and limitation of the number of sample group at 11 years (n=132), 9 years (n=85) and with 7-8 years and 10 years having almost similar number of population as shown in Table 3, thus maybe making the comparison in this sample group statistically non-significant.

For classifying based on gender as shown in Table 8, it can be seen that the sample group consisting of different genders had scores from TONI-3 having no significant difference at a statistical level of .05, which was consistent with the results of the study made by the originators of the test namely Brown, Sherbenou, and Johnsen (31), who reported that different genders had scores that were not different when TONI-3 was used.
Also, when comparison was made on the difference of the scores from TONI-3 which were classified according to age, as shown in Table 9, there was significant difference at statistical level of .001, in the average scores in the age groups. The average scores showed tendency to increase when age also increased except in age group of 11:00-11:11 which had average scores slightly lower than 10:00-10:11 but when testing was done on the difference between age pairings, as shown in Table 10, it was found that group 1 in particular (ages of 7:00-7:11) only that had significant differences at a statistical level of .05) with other groups. Looking at the average scores, however, showed that they tended to increase when age increased which indicated that the visual perception, relationship dimension and ability for abstract thinking of the Thai child had developed based on age. This was consistent with the theory on intellectual development by Piaget who stated that the intellectual development on knowledge and understanding is based on seniority and in order (16), in a gradual manner. The development increases ever since infancy to teen-age. The sample group in this study consisted of children aged 7 years to 11 years and 11 months and these ages belonged to the Concrete Operational Period when a child understands about whatever that completely supports the past and recognizes the use of relative thinking, arranging what goes before and what comes after and thinking backwards, more of cause and effect (16) which supports the study of standardization of TONI-3 test (22). This showed that the ability to think and solve problems with abstractions, increased depending on age that also increased likewise. The results of this study showed a slight movement from the initial theory, which meant that based on the theory, during the age period of 11:00-11:11, the average score must be the highest. When considering the movement, it might have come from the selection of the sample group, which was shown that most children (132) were at ages 11 years to 11 years and 11 months (usually at the 6th grade) or about 32 percent of the sample population. At the same time, the sample group with ages of 10 years to 10 years and 11 months consisted of 59 children. As seen from the standard deviation value, it was shown that the sample group had a standard deviation of 10.21 which was a high value when compared with other age groups. From this reason, it might lead to the average score of the children in this sample group to be much lower than 10 years to 10 years and 11 months.
Analysis of difference between group testing and individual testing using the TONI-3 and CPM tests

The results of the classification of IQ level of TONI-3 in both group and individual testing, as shown in Table 6, indicated that when doing the individual testing, this would result to the classification that was different from group testing. In group testing, the classification resulted to 7 levels but in individual testing, the classification has led to only 5 levels where no child was classified as having an IQ score of Borderline and Extremely Low at all. When the group test was conducted, there were some children who were classified as having IQ at the Borderline and Extremely Low at 6.7 and 6.7%, respectively. It can be seen that when TONI-3 was used in a group test for the sample group with ages of 9 years and 9 years and 11 months, the percentage of the sample group having an IQ level at Low Average was higher than the average. This was due mainly to the sample group being much younger thus not being able to understand the questions much clearly or not understanding the method of answering the questions. But when the test was given individually, the one administering the test was able to observe the behavior when the child worried or could not understand and the person administering the test could immediately make an example for the child to see thus enabling the latter to understand the question and the method much better resulting to an increase in the IQ level much higher than in the group test.

When the CPM test was administered, it was found that individual testing was able to classify the sample group into 5 groups likewise (as shown in Table 6) but the IQ levels were different from TONI-3 in two levels. TONI-3 test was able to classify the group to the highest level of Very Superior and the lowest level of Low Average while CPM was only able to classify the highest level at Superior and lowest level at Borderline. Thus it can be seen that the intelligence level of the child from the individual testing was increased similarly.

When the test was administered in both group and individual manner by TONI-3 test, it was found that the IQ levels of the sample group were different due to TONI-3 was created with the purpose of using it as an individual test only in order to
observe the behavior of the child as in facial expression and other bearings whether there was understanding in the method of the test or not, aside from the ones explained by the person giving the test. The test was stopped when Ceiling Age of the child being tested was reached as indicated by the test manual. Thus in using TONI-3 test, there is a need for an individual test to be implemented to attain correct results and follow the objectives of the creator of the test and the results of the individual test from the sample group would have the IQ levels much higher than from the group test. On the part of the CPM test administered either in group or individually, it was found that IQ levels were not significantly different at a statistical value of .05. This could be due to the characteristics of the test which had no point of non-verbal communication between the children being tested and the person administering the test in a detailed manner as done in TONI-3 test.

In this study, it was found that TONI-3 test was able to screen intelligence of the Thai children but must be implemented in an individual manner so the results would be more correct than the limitations and obstacles that occur during data collection as indicated below.

In the implementation of the test in a group manner, the results might be dispersed because of the various causes such as: every time the group test was implemented took a longer time of 50-60 minutes and the small children with ages from 7 to 9 years may lack the concentration to be interested or there might be competition in the sample group to finish the test, there might be cheating or looking at the other’s answers, talking while taking the test, and others. These might affect the test to be less credible so in future studies, an individual test must be implemented as indicated in the test manual in order to prevent those above mentioned problems because in the implementation of the individual test, the time required to finish the test is only 15-20 minutes thus not allowing children to become bored of the test and by nature of the test, there is nonverbal communication between the children being tested and the person administering the test which indicates the understanding of the abstract thinking in answering the questions. Aside from these, the person administering the test could observe the behavior of the person being tested. When felt and observed that the latter becomes tired, he is allowed to take a break during the test in order to
decrease the tiredness in answering the test most particularly among children with ages between 7 and 9 years.

The care taken particularly on the characteristics of the sample group could be another factor that might have caused the dispersion of the results of the study. In the sample, some children knew that this was an intelligence test so they wanted to have higher scores so they tried to look at the answers of their more intelligent friends or other children used favoritism or special treatment during the entire course of the study. Other children were not fully attentive to the test but could not refuse to take the test. These caused the test scores to be less reflective of the actual scores of some children. Thus there should be a control of the sample group to be more prepared to take the test by establishing good relations with the sample group or more rules should be established to allow children to be more attentive in taking the test. These could be done characteristically when the test is administered individually because evaluation of the work is done throughout the course of the test and solving the problem not as when the person taking the test takes the test by himself until finishing it. The difference between the results of the individual and the group tests shows the comparison as seen from this study.
CHAPTER V

CONCLUSION AND RECOMMENDATION

This research was conducted in order to study the properties of measurement of TONI-3 in terms of its various validities such as in this particular study, the criterion-related validity which involved the determination of correlation between scores from TONI-3 and CPM, and construct validity in terms of age differentiation and practicality based on their application and limitations in administering the test.

The sample group used in this study consisted of children studying in the primary school and whose ages ranged from 7 years to 11 years and 11 months. A total of 380 children, 172 male and 208 female, came from Wat Amarintharam School in Zone 3 of Bangkok Metropolis under the Office of the Commission on Basic Education, Ministry of Education. Data collection was started with the document of acceptance given to the parents of the children in the sample group to sign their approval after which data gathering was done according to the time table and with the aid of a research assistant who had agreed to understand the method and the steps in data collection. In the initial part of data collection, the test was administered in group testing using a time period of 50-60 minutes. In the group of younger school children studying in grades 1, 2 and 3, not more than 5 children were allocated for one test administrator, while for older children in grades 4, 5 and 6, not more than 10 children were allocated per one test administrator. Meanwhile, in the second part of data collection, the test was given to individuals from the sample group but only to those 60 individuals whose ages range from 9 years to 9 years and 11 months. The test was administered in a period of 40-50 minutes. Prior to the administration of the two tests, the test administrator built a good rapport with the sample group after which, explained the objectives of the test and the method of implementing the test to the sample group for their understanding. Then the tests were administered starting with the TONI-3 test and upon its completion, was followed by the CPM test.
Conclusion

In this research, the main objectives were to study the properties of the TONI-3 test in terms of its validity such as its criterion-related validity by determining the relationship between the scores of the person taking both tests, construct validity through age differentiation and practicality in the point of view of its benefits and limitations in using the test. These could be summarized as follow:

1. From the examination of the validity of TONI-3 test in two aspects, criterion-related validity in terms of the relationship between test scores and construct validity in terms of age differentiation, it was found that TONI-3 had a linear correlation with CPM at a moderate level at a statistical significance (.001) with a Pearson’s moment correlation coefficient of .466.

2. When considering the fitness in the classification of the intelligence level between TONI-3 and CPM from the group testing, it was found that the fitness was less at only 35 percent with an intelligence level of 90-109 (Average) having 66.3 percent fitness. This was followed by intelligence level at 110-119, 80-89 and lower than 70, respectively. There were 3 stages of the intelligence level which were found to have test scores from both TONI-3 and CPM that showed no fitness and these included 70-79, 120-129 and 130 and above.

3. For testing on the difference of scores of the sample group which consisted of male and female respondents who were administered the TONI-3 test, it was found that the sample group consisting of different genders obtained test scores that were not significantly different (statistical level of .05) and from testing of sample group belonging to different ages, only the sample group with ages of 7 years-7 years and 11 months had intelligence scores that were different from other groups with significant difference (statistical value of .001) with each group. When comparison was made of their average scores, results showed that when age of the sample group was increased, the average scores also increased.

4. Based on the comparison between the group and individual tests of the two intelligence tests, it was found that TONI-3 test gave results or scores that were significantly different (statistical value of .05) with sample group that were
individually administered of the test having higher scores than those tested in group. On the part of the CPM test, there was no significant difference (statistical value of .05). When comparison was made on the classification, results showed that when TONI-3 was used in a group manner, it was able to classify the sample group into 7 levels. However, even though group test is able to classify more the intelligence level of the sample group than the individual test, it was still found that in group test, there was still a sample group with IQ level on the Borderline and Extremely Low at 13.4 percent more than the norm. But when individual test was administered, it was able to classify to only 5 levels without sample group having IQ level at Borderline and Extremely Low.

**Recommendations**

The recommendations in using the results of this study are explained as follow:

1. In conducting the TONI-3 test to screen out the intelligence ability, it must be implemented according to the test manual only which meant that the test must be administered in an individual manner because in the method of testing, there must use altogether observations the various expressions or bearings in order to find out whether the person taking the test fully understands the method of testing and must stop when one reaches the Ceiling Age of the one being tested. If administered as a group test, these details could not be according to the correct method in testing. Because of those reasons, it becomes inappropriate to implement the test in group manner as it makes the results of the measurement dispersed from the actual value.

2. In the study about the TONI-3 intelligence test in the future, it may perhaps be done in a computerized manner in order to improve the convenience and interest of the test but this should still be done strictly according to the rules included in the test manual. The persons administering the test must altogether use observations of the persons being tested in the initial part of the test, whether they have actual understanding of the method of implementing the test thus being able to allow the latter to take the test by themselves.
3. In the implementation of the TONI-3 and CPM tests for use in the screening of the intelligence abilities, there are two points of observation. If the person administering the test wants to screen out the intelligence into different levels, he must use TONI-3 test because this test is able to classify them into extreme levels much better than CPM test together with its better ability to classify the group into irregular or abnormal levels but if one needs to screen intelligence levels into normal levels, it is suggested that CPM test should be used.

4. CPM test should not be applied for use in screening out the intelligence using children whose ages range from 9-11 years because it is too difficult for them as CPM test is a Ceiling Age for these children.


3. โยวิน ศิริธรรมและคณะ. วิจัยทางวิทยาศาสตร์. กรุงเทพฯ: สภากาชาดไทย; 2533.

4. สมพร สุขศิลป์. แบบทดสอบทางจิตวิทยา. กรุงเทพมหานคร: สภากาชาดแห่งจุฬาลงกรณ์มหาวิทยาลัย; 2545.


7. สุชีรา ภัทรายุตวรรตน์, กนกรัตน์สุขะตุงคะ, จริยาจันตระ, วิลาสินีชัยสิทธิ์, กิจิ บรรณกุลโรจน์และสร้อยสุดาอิ่มอรุณรักษ์. การสร้างเกณฑ์ปกติของแบบทดสอบ Coloured Progressive Matrices ในกลุ่มประชากรไทย. วารสารจิตวิทยาคลินิก; 34(1)(มกราคม), 58-70; 2546.


25. เชิดศักดิ์ โฆวาสินธุ, การวิจัยทางการศึกษา. กรุงเทพมหานคร: สำนักพิมพ์โอเดียนสตรี; 2549.

26. ธารา ทวายเทาน, การศึกษาความเพื่องรัฐวิทยาปัจจุบันของแบบทดสอบวัดความรู้, สำหรับนักเรียนขั้นที่ 3, ปีการศึกษา 2549, วิทยานิพนธ์ปริญญาการศึกษา. มหาวิทยาลัยบัณฑิต, 2545.

27. ส่งเสริม เจาดีศิลป์, ยุทธการทดสอบทางจิตวิทยา. พิพิธภัณฑ์ที่ 2, เชียงใหม่: เชียงใหม่ 2543.

28. วิโรจน์ ศิริพาณิชย์, หลักการสร้างแบบสอบถามทางจิตวิทยาและทางการศึกษา. กรุงเทพฯ: สาขาจิตวิทยา, คณะศิลปศาสตร์, มหาวิทยาลัยธรรมศาสตร์; 2533.

29. สุทธิชัย วชิรพานิช, หลักการวิจัยและการใช้สถิติในการวิจัยจิตวิทยา. ภาควิชาจิตวิทยาศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, 2533.


34. กนกรัตน์ ทุ่งนุ่ง, ดุสิตจิตวิทยาคลินิก. กรุงเทพฯ: แมคคลีมี่เดล_values; 2538.


38. กนกรัตน์ สุขะตุงคะ และคณะ. การพัฒนาเกณฑ์ปกติคะแนนมาตรฐานเชาวน์ปัญญาในเด็กไทยจากแบบทดสอบเชาวน์ปัญญาของ Raven (Coloured Progressive Matrices – CPM) ในกลุ่มประชากรไทย. วารสารสมาคมนักจิตวิทยาคลินิก. 37(2549) : 10-22.


40. กีรติ บรรณกุลโรจน์. การพัฒนานิพนธ์แบบทดสอบ Visual-Motor Integration 3rd revision (VMI 3R) ในกลุ่มนักเรียนประถมศึกษา ในเขตกรุงเทพมหานคร. [วิทยานิพนธ์ปริญญา มหาบัณฑิต]. กรุงเทพฯ : นิติศาสตร์ มหาวิทยาลัยมหิดล; 2542.
**BIOGRAPHY**

<table>
<thead>
<tr>
<th>NAME</th>
<th>Miss Woraya Sroythong</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATE OF BIRTH</td>
<td>1 October 1983</td>
</tr>
<tr>
<td>PLACE OF BIRTH</td>
<td>Bangkok, Thailand</td>
</tr>
<tr>
<td>INSTITUTIONS ATTENDED</td>
<td>Chiang Mai University, 2006:</td>
</tr>
<tr>
<td></td>
<td>Bachelor of Science (Psychology)</td>
</tr>
<tr>
<td></td>
<td>Mahidol University, 2008:</td>
</tr>
<tr>
<td></td>
<td>Master of Science</td>
</tr>
<tr>
<td></td>
<td>(Clinical Psychology)</td>
</tr>
<tr>
<td>SCHOLARSHIP</td>
<td>Scholarship supported from</td>
</tr>
<tr>
<td></td>
<td>Siriraj Graduate Scholarship</td>
</tr>
</tbody>
</table>