

บทที่ 12

แบบทดสอบทางสติปัญญา (Intelligence Test)

จุดประสงค์การเรียนรู้

1. อธิบายทฤษฎีโครงสร้างทางสติปัญญาได้
2. เปรียบเทียบความแตกต่างระหว่างคำว่า “สติปัญญา” กับ “ความถนัด” ได้
3. อธิบายลักษณะของแบบทดสอบวัดเชาวน์ปัญญาของ Stanford-Binet ได้
4. อธิบายวิธีการทำแบบทดสอบเชาวน์ปัญญาของ Wechsler ทั้ง 3 ประเภทได้
5. อธิบายลักษณะของข้อสอบ TONI ได้
6. ยกตัวอย่างแบบทดสอบความถนัดที่ใช้กันแพร่หลายได้

คำโครงเนื้อหา

1. ทฤษฎีโครงสร้างทางสติปัญญา
2. ความแตกต่างระหว่างคำว่า “สติปัญญา” กับคำว่า “ความถนัด”
3. Stanford-Binet Intelligence Test
4. Wechsler Intelligence Test
5. TONI (Test of Nonverbal Intelligence)
6. แบบทดสอบความถนัด

เนื้อหาสาระสำคัญ

ความหมายของคำว่า “intelligence” นั้น ปกติอาจหมายถึงสมรรถภาพทางด้านต่าง ๆ ดังนี้

1. การคิดแบบสอบถาม
 2. การเรียน
 3. การนำประสบการณ์ใหม่มาประสมประสานเข้ากับสถานการณ์ใหม่
- นักจิตวิทยาหลายท่านได้ให้คำจำกัดความของคำว่า Intelligence ไว้ดังนี้

Binet กล่าวว่า “เป็นสมรรถภาพในการตัดสินใจ ในการให้เหตุผลและความเข้าใจในเรื่องต่าง ๆ ได้เป็นอย่างดี”

Goddard กล่าวว่า “เป็นอัตราความสามารถของแต่ละบุคคลที่จะแก้ไขปัญหาเฉพาะหน้าและคาดการณ์ล่วงหน้าได้”

Terman กล่าวว่า “เป็นความสามารถที่จะคิดออกมาในแบบของนามธรรม”

1. ทฤษฎีทางโครงสร้างของสติปัญญา (Theories of Intelligence Structure)

ความเคลื่อนไหวในการวัดทางสติปัญญาที่ได้เริ่มมีขึ้นในศตวรรษที่ 19 นี้เอง Sir Francis Galton ซึ่งเป็นนักชีววิทยาชาวอังกฤษ เป็นผู้ริเริ่มนำหลักของการแปรผันและการคัดเลือกโดยศึกษาต่อมาจากญาติของเขา คือ Darwin โดย Galton ต้องการจะศึกษาในแง่ของการวัดคุณลักษณะของกรรมพันธุ์ ในด้านรูปร่างและคุณลักษณะของจิตใจ Galton จึงสร้างเครื่องมือขึ้นหลายชนิดเพื่อจะวัดสิ่งเหล่านี้ เขาเชื่อว่าการทดสอบความแตกต่างทางด้านประสาทสัมผัสและการได้กลับของผู้ถูกทดสอบ จะเป็นตัวที่อาจนำมาใช้ประมาณค่าขององค์ประกอบทางสติปัญญาได้ ดังนั้น แบบทดสอบของ Galton จึงมีมากมายหลายแบบ ต่อมา James Mc. Keen Cattell นักจิตวิทยาชาวอเมริกัน ก็ได้ตั้งทฤษฎีว่า ความเฉียบแหลมของประสาทสัมผัส ความรวดเร็วในการตอบสนอง หรือสิ่งที่คล้าย ๆ กันนี้ จะแสดงให้เห็นถึงองค์ประกอบของสติปัญญา Cattell เป็นผู้นำคำว่า “mental test” มาใช้เป็นคนแรกในปี 1890 แบบทดสอบของเขาใช้สำหรับวัดความแข็งแรงของกล้ามเนื้อ ความรวดเร็วในการเคลื่อนไหว ความอ่อนไหวต่อความเจ็บปวด ความสามารถในการจำแนกน้ำหนักร

นักจิตวิทยาที่ดำเนินการตาม Galton และ Cattell โดยวัดสิ่งที่คล้าย ๆ กันนี้ก็คือ Jastrow ส่วน Binet และ Henry ได้เริ่มทำการศึกษาวิจัย โดยการวัดคุณลักษณะต่าง ๆ เป็นต้นว่า ความจำ ความเอาใจใส่ ความเข้าใจ และเริ่มสนใจที่จะวัดสิ่งที่ซับซ้อนมากขึ้นไปอีก ต่อมาปี 1905 Binet และ Simon จึงได้ปรับปรุงแบบทดสอบ Individual Intelligence (The Binet Scale) ขึ้นโดย

ออกแบบให้แบบทดสอบนี้สามารถวัดเชาวน์ปัญญาได้หลาย ๆ ระดับ

ในปี 1927 Spearman ก็ได้ก่อตั้งทฤษฎี 2 องค์ประกอบ (Two-factor theory) ขึ้นมา คือเขาให้ความหมายของทฤษฎีนี้ว่า สถิติปัญญาประกอบด้วยองค์ประกอบ 2 อย่าง คือ

1. G factor
2. S factor

G factor หมายถึงองค์ประกอบทั่ว ๆ ไปที่ทุกคนควรจะมี ส่วน S-factor นั้นเป็นความสามารถพิเศษ อาจจะมีตั้งแต่ $S_1, S_2 \dots S_n$.

ต่อมา Thurstone ใช้วิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) โดยปรับปรุงจากทฤษฎีของตัวประกอบพหุคูณออกมา เป็นแบบทดสอบที่วัดความสามารถเบื้องต้นของสติปัญญา Vernon ได้แบ่งโครงสร้างของความสามารถออกเป็นกลุ่มใหญ่ ๆ 2 กลุ่ม คือ ความสามารถในการเรียนภาษา และความสามารถในการใช้เครื่องมือต่าง ๆ ทฤษฎีของ Vernon ไม่กว้างขวางเหมือนของ Spearman แต่ก็กว้างกว่าทฤษฎีองค์ประกอบของ Thurstone

Guilford ได้เสนอโครงสร้างทางสติปัญญา ออกมาเป็น 3 ประเภท คือ

- | | |
|--------------------------|-----------|
| 1. กระบวนการ (Operation) | มี 5 ชนิด |
| 2. ผลิตผล (Product) | มี 6 ชนิด |
| 3. เนื้อหา (Content) | มี 4 ชนิด |

ฉะนั้น โครงสร้างทางสติปัญญาตามแนวของ Guilford จึงประกอบด้วย $4 \times 5 \times 6 = 120$ cells

จากที่กล่าวมาแล้วจะเห็นได้ว่า ทฤษฎีโครงสร้างทางด้านสติปัญญามีแตกต่างกันออกไปหลายทฤษฎี เพราะฉะนั้นแบบทดสอบวัดเชาวน์ปัญญา จึงมีมากมายหลายฉบับแตกต่างกันออกไป สิ่งสำคัญของผู้นำแบบทดสอบไปใช้ก็คือจะต้องเข้าใจด้วยว่า ผู้สร้างแบบทดสอบได้ให้คำนิยามแบบทดสอบนั้น ๆ ว่าจะวัดอะไรได้บ้าง

2. ความแตกต่างระหว่างคำว่า Intelligence กับ Aptitude

คำว่าเชาวน์ปัญญาและความถนัดนั้น ส่วนมากจะใช้สลับเปลี่ยนกันอยู่เสมอ แท้ที่จริงแล้วมีความหมายแตกต่างกันดังนี้ คือ

1. แบบทดสอบวัดเชาวน์ปัญญานั้น ใช้เพื่อวัดความสามารถทั่ว ๆ ไป แต่ถ้าต้องการจะวัดความสามารถเฉพาะอย่างใดอย่างหนึ่ง หรือวัดความสามารถหลาย ๆ ด้าน เราก็เรียกว่าเป็นแบบทดสอบความถนัด เช่น ข้อสอบของ Wechsler Adult Intelligence Scale ซึ่งจะวัดองค์ประกอบหลาย ๆ ด้าน จัดเป็นแบบทดสอบวัดเชาวน์ปัญญา

2. ความแตกต่างในข้อนี้ก็เนื่องมาจากประวัติความเป็นมาของแบบทดสอบทั้ง 2 ประเภท ในสมัยเริ่มมีแบบทดสอบวัดเชาวน์ปัญญาขึ้นมา นั้น นักจิตวิทยาเข้าใจว่าเชาวน์ปัญญาของเด็กนั้นเป็นสิ่งติดตัวเด็กมาตั้งแต่เกิด ไม่อาจจะเปลี่ยนได้ ความเข้าใจเรื่องนี้ นับว่ายังไม่ถูกต้อง ดังนั้น เพื่อจะหลีกเลี่ยงความหมายที่ว่าเชาวน์ปัญญาเปลี่ยนแปลงไม่ได้ ผู้สร้างแบบทดสอบหลาย ๆ คน จึงนิยมใช้คำว่า ความถนัด (aptitude) มากกว่า เพราะว่าแบบทดสอบความถนัด นับว่ามีประโยชน์อย่างยิ่งที่จะใช้ทำนายความสำเร็จในการเรียนได้ และบางคนก็เชื่ออย่างฝังใจว่าแบบทดสอบความถนัดทางการเรียนจะเป็นสิ่งพยากรณ์ได้แม่นยำที่สุด

3. ในการสร้างแบบทดสอบเพื่อวัดเชาวน์ปัญญาของคนโดยหลีกเลี่ยงอิทธิพลทางด้านวัฒนธรรม และการศึกษา คนส่วนมากที่ต้องการเช่นนั้น จึงหันมานิยมใช้แบบทดสอบวัดเชาวน์ปัญญามากกว่า ซึ่งก็หมายความว่าผู้นิยมใช้มักจะมีแนวโน้มว่าสติปัญญาของคนเปลี่ยนไม่ได้

แบบทดสอบเชาวน์ปัญญา ที่จะแนะนำให้รู้จัก คือ

1. Stanford - Binet เป็นแบบทดสอบวัดเชาวน์ปัญญาที่ใช้สอบทีละคน
2. Wechsler Scale แบ่งเป็น
 - 2.1 Wechsler Intelligence Scale for Children (WISC)
 - 2.2 Wechsler Adult Intelligence Scale (WAIS)
 - 2.3 Wechsler Pre - school and Primary Scale of Intelligence (WPPSI)
3. TONI (Test of Nonverbal Intelligence)

3. แบบทดสอบวัดเชาวน์ปัญญาของ Stanford - Binet



ในศตวรรษที่ 20 Binet ได้สร้างแบบทดสอบ I.Q. โดยทำเป็น batteries of Tests มีทั้งหมด 10 subtests คือ

1. Memory เป็นการวัดเกี่ยวกับความจำทั่วไป โดยมีภาระสิ่งหนึ่งกระตุ้นให้เกิดผลสะท้อนออกมา แบ่งออกเป็น

- ก. การใช้สายตาจำรูปทรงเรขาคณิต
- ข. จำประโยค
- ค. จำโน้ตดนตรี
- ง. จำสีต่างๆ
- จ. จำตัวเลข

2. Images ให้บรรยายภาพที่เห็นอย่างชัดเจนโดยกำหนดตัวกระตุ้นมาให้

3. Imagination แบ่งเป็น

- ก. ink - blot
- ข. คำที่เป็นนามธรรม
- ค. ให้แต่งเรื่องหรือเติมเรื่องจากภาพ
- ง. ให้สร้างวลีให้มากที่สุดเท่าที่จะทำได้ โดยให้มีคำนามที่กำหนดให้อยู่ด้วย

วลีละ 3 คำ

จ. แต่งเรื่องตามหัวข้อที่กำหนด

4. Attention

- ก. จำเรื่องที่กำหนดให้หลาย ๆ เรื่อง
- ข. ให้ขีดและนับจำนวนตัวที่มีอักษร ในหน้าที่กำหนดให้
- ค. ให้นับจำนวนที่นาฬิกา 2 เรือนตีบอกเวลาใน Speed ต่าง ๆ กัน
- ง. ให้ทำตามคำสั่งในทันทีที่ได้รับคำสั่ง

5. Comprehension

ก. ให้อธิบายความเคลื่อนไหวของจักรเย็บผ้าหรือแมลงที่เคลื่อนไหวได้โดยกำหนดเวลาให้

- ข. บอกความคล้ายคลึงและความแตกต่างของคำ Synonyms
- ค. วิเคราะห์ประโยคที่ไม่ถูกต้องได้

6. Suggestibility

ก. แสดงให้เห็นถึงความแตกต่างของระยะทาง และค้นหาที่มาของกลิ่นต่างๆ เพื่อจะวัดดูว่าจะยอมให้คนอื่นมาแนะนำหรือไม่

- ข. คิดคำนึงออกมาโดยการคาดการณ์ล่วงหน้า
- ค. กำหนดอารมณ์ได้
- ง. ไม่ยอมทำตามที่คนอื่นสั่ง

7. Asthetic Appreciation

- ก. เปรียบเทียบสีต่าง ๆ ที่จิตรกรต่างคนต่างทำได้
- ข. รู้จักความสัมพันธ์ของสี
- ค. เข้าใจเนื้อเพลง

8. Moral sentiments สัมผัสผลสะท้อนจากการดูภาพในแง่มุมต่าง ๆ

9. Mascular force and force of will

- ยกน้ำหนัก เครื่องทดสอบพลังต่าง ๆ และกำลังใจ

10. Motor Skill and visual judgement

- ก. ร้อยเข็มเย็บผ้า
- ข. เกาะชิงหระ
- ค. ประมาณได้ว่าจะต้องต่อเส้นตรงก็เส้นจึงจะได้เท่ากับเส้นที่กำหนดไว้

ต่อมา Stoddard ได้เอา test ของ Binet ไปแปลงใหม่ แต่ก็ไม่แพร่หลาย Terman จึงเอามาทำต่ออีก 10 ปี จึงมีชื่อเสียงมาก เรียกว่า Standford of Binet Intelligence Test

4. Wechsler Intelligence Test มี 3 ประเภท คือ

4.1 Wechsler Intelligence Scale for Children (WISC)

แบบทดสอบนี้ใช้กับเด็กตั้งแต่อายุ 5 ขวบขึ้นไป โดยใช้ทฤษฎีของ Burt & Vernon ที่กล่าวถึง fluid ability กับ crystalize ability

ข้อจำกัดบางประการในการใช้ข้อสอบ WISC

1. ใช้กับเด็กอายุระหว่าง 5 ถึง 15 ขวบ
2. ใช้กับเด็กที่มี I.Q. ประมาณ 46 ถึง 154
3. การให้คะแนน Comprehension, Similarities และ Vocaburary ให้ยากมาก
4. ใช้กับเด็กในกลุ่มที่จำกัดเชื้อชาติ
5. คะแนนที่ได้แปลผลยาก

แบบทดสอบ WISC แบ่งเป็น 2 ส่วน คือ

1. Verbal Scale ประกอบด้วย
 - 1.1 General Information
 - 1.2 General Comprehension

1.3 Arithmetic

1.4 Similarities

1.5 Vocabulary (Digit Span)

2. Performance Scale ประกอบด้วย

2.1 Picture Completion

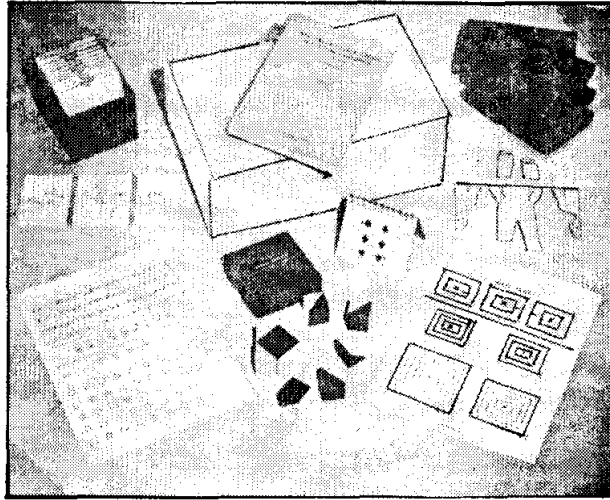
2.2 Picture Arrangement

2.3 Block Design

2.4 Object Assembly

2.5 Coding (or Mazes)

แบบทดสอบ WISC ได้ปรับปรุงขึ้นมาเพื่อต่อต้านกับ age - scale ของ Stanford - Binet เนื่องจากการสอบเป็นชุด ๆ นี้ใช้เวลาสั้น จึงทำให้เป็นที่สนใจแก่เด็ก สำหรับพวก Performance นั้น เวลาให้คะแนนก็ดูเวลาและความถูกต้อง เมื่อสอบครบ 12 ชุด (คือใช้ข้อสอบสลับกันระหว่าง Coding และ mazes และ Digit Span กับ Vocabulary) ก็เอาคะแนนมารวมแยกแต่ละชุด ก็สามารถจะคำนวณหาเฉลยของเด็กออกมาได้



เกณฑ์ปกติและการให้คะแนน

การให้คะแนนของ WISC คะแนนดิบที่ได้จะเปลี่ยนเป็นคะแนนมาตรฐานและเปรียบเทียบในกลุ่มเดียวกัน มีตารางคะแนนสำหรับเปรียบเทียบในช่วง 4 เดือน ตั้งแต่อายุ 5 ปี ถึง 15 ปี กลุ่มตัวอย่างมาตรฐานของ WISC มีเด็กชาย 100 คน เด็กหญิง 100 คน ที่มีอายุต่างระดับกันตั้งแต่ 5 ปี ถึง 15 ปี ซึ่งให้ผลรวมกันถึง 2,200 อย่าง

ค่าความเชื่อมั่น (Reliability)

หาโดยวิธี Split - half จากรายงานของ WISC คะแนนจาก Verbal Performance และ Full Scale ให้ค่าความเชื่อมั่นที่พอเชื่อถือได้ ค่าความเชื่อมั่นนี้คำนวณแยกกันจากกลุ่มตัวอย่างที่มีอายุ 7½ ปี, 10½ ปี และ 13½ ปี แต่ละกลุ่มประกอบด้วย 200 ตัวอย่าง เทคนิคแบบข้อคู่ข้อคี่ไม่ได้นำมาใช้กับข้อสอบแบบ Coding & Digit Span เพราะคะแนนจากแบบทดสอบทั้ง 2 ชุดนี้สัมพันธ์กัน ค่าความเชื่อมั่นของการสอบเด็กทั้งหมดมีค่าตามระดับอายุเท่ากับ .92, .95 และ .94 ตามลำดับของ Verbal ได้ .88, .96 และ .96 สำหรับ Performance Scale ได้ .96, .89 และ .90 ทั้งแบบทดสอบรวมทั้งหมด แบบทดสอบ Verbal และแบบทดสอบ Performance มีค่าความเชื่อมั่นที่พอจะให้ประโยชน์ตามความมุ่งหมายของการทดสอบมากที่สุด ส่วนค่าความคลาดเคลื่อนมาตรฐาน (Standard error of measurement) มีพิสัยอยู่ระหว่าง 3-5.61

หลังจากนั้น 4 ปี WISC ก็มีความเชื่อมั่นพอ ๆ กับของ Stanford Binet จากการทดสอบเด็กเกรด 5 จำนวน 60 คน ซึ่งมาทดสอบครั้งหลังเมื่อเด็กอยู่เกรด 9 พบว่า ค่าสหสัมพันธ์ระหว่างแบบ Stanford - Binet กับ WISC มีค่าเท่ากับ .78 ส่วน Full scale, Verbal และ Performance มีค่าสหสัมพันธ์กันเท่ากับ .77, .77 และ .74 ตามลำดับ

ความเที่ยงตรง (Validity)

ในคู่มือของ WISC ไม่มีการกล่าวถึงเรื่องความเที่ยงตรง แต่มีตารางเทียบคะแนนมาตรฐานกับคะแนนจากแบบทดสอบที่อายุระดับต่าง ๆ กัน แต่ก็ไม่ได้ประเมินผลของข้อมูลกับเกณฑ์ที่วางไว้

การประเมินค่าทั่วไป

แบบทดสอบ WISC เมื่อเปรียบเทียบกับสเกลวัดเชาว์อื่น ๆ แล้ว จะเห็นว่ามีคุณภาพด้านการสร้างขนาดและตัวแทนของกลุ่มตัวอย่างที่จะใช้ทำ norms การปรับปรุงข้อสอบให้เป็นมาตรฐานนั้นถือว่า มีความเชื่อถือได้พอเพียง แต่มีสเกลบางอย่างได้ข้อมูลที่ขาดความเที่ยงตรง ซึ่งเป็นจุดอ่อนที่สำคัญของแบบทดสอบนี้

4.2 Wechsler Adult Intelligence Scale (WAIS)

Wechsler Scale แบบแรกที่ถูกจัดกันนั้นชื่อ Wechsler Bellevue Intelligence Scale ซึ่งเริ่มพิมพ์ในปี ค.ศ. 1939 โดยมีจุดมุ่งหมายที่จะวัดเชาว์ผู้ใหญ่

WAIS ได้พิมพ์ขึ้นปี 1955 ประกอบด้วย 11 subtests ถือว่าเป็น Verbal Scale 6 subtests เป็น Performance Scale 5 subtests ดังนี้

Verbal Scale

1. *Information* มีคำถาม 29 คำถาม โดยถามครอบคลุมเรื่องราวหลายอย่างที่ผู้ใหญ่มีอยู่หรือได้รับมาจากวัฒนธรรม ทั้งนี้ ก็พยายามหลีกเลี่ยงความรู้ด้านวิชาการ หรือความรู้พิเศษ และอาจรวมคำถามเกี่ยวกับเรื่องทั่วไป เป็นการตรวจสอบทางจิตในการตั้งเกณฑ์ระดับเชาวน์สำหรับแต่ละคน และการปฐมนิเทศภาคปฏิบัติ

2. *Comprehension* มีคำถาม 14 ข้อ แต่ละข้อเป็นการอธิบายถึงว่าจะทำอะไรภายใต้กรณีนั้น และทำไมต้องปฏิบัติไปอย่างนั้น มีความหมายของสุภาษิตและอื่น ๆ อยู่ด้วย เหมือนการถามความเข้าใจของ Stanford Binet แต่มีเนื้อหาเฉพาะเจาะจงเลือกให้คล้องจองกับความสนใจ และเป็นกิจกรรมของผู้ใหญ่

3. *Arithmetic* มีปัญหา 14 ข้อ คล้าย ๆ กับเลขปัญหาในโรงเรียนประถมศึกษา แต่ละปัญหาจะอ่านให้ฟัง แล้วให้แก้ปัญหานั้นโดยไม่ต้องใช้ดินสอและกระดาษ

4. *Similarities* มีคำถาม 13 ข้อ เป็นการเปรียบเทียบของสองสิ่งที่เหมือนกัน

5. *Digit Span* คือ การออกเสียงซ้ำตามที่กรรมการพูดให้ฟัง กรรมการจะพูดออกเสียงตัวเลข 3 ถึง 9 ตัว ในตอนที่ 2 ให้ว่าตามจากเลข 2 – 8 ตัว แต่พูดถอยหลัง

6. *Vocabulary* มีคำศัพท์ 40 คำ เพิ่มความยากขึ้นตามลำดับ ใช้ทั้งวิธีอ่านออกเสียงให้ฟังและเขียนให้ดู แล้วถามความหมายของคำแต่ละคำ

Performance Scale

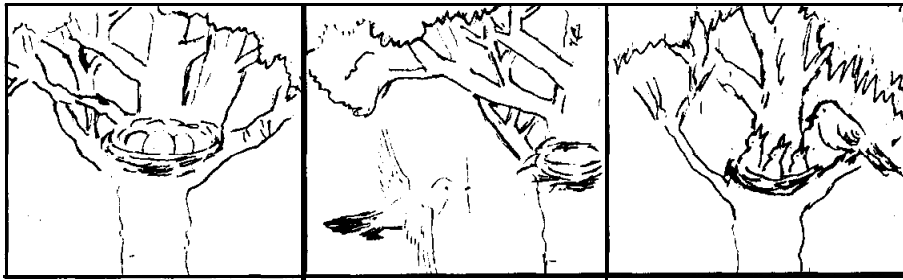
1. *Digit Symbol* เป็นแบบทดสอบที่ให้ผู้สอบใส่ตัวอักษรที่คุ้นเคยลงในช่อง เป็นแบบไม่ใช้ภาษา คำตอบประกอบด้วยสัญลักษณ์ 9 ตัว จัดคู่กับตัวเลข 9 ตัว คำตอบอยู่ข้างหน้า ผู้สอบต้องตอบโดยใส่สัญลักษณ์ให้มากที่สุดเท่าที่จะทำได้ ใต้จำนวนบนกระดาษคำตอบให้เวลา 1½ นาที

2. *Picture Completion* มีบัตร 21 แผ่น แต่ละแผ่นมีภาพ 1 ภาพนั้น มีบางส่วนของภาพอยู่ การสอบก็ให้บอกว่า ภาพแต่ละภาพขาดอะไรอยู่

3. *Block Design* แบบทดสอบนี้ Wechsler ได้ดัดแปลงมาจาก Rehs Block Design Test โดยทำเฉพาะด้านสีแดง สีขาว และสีแดงขาวเท่านั้น ผู้สอบจะต้องถอดแบบสลับซับซ้อนนั้น จาก 4 – 9 ชั้น

4. *Picture Arrangement* ข้อคำถามแต่ละข้อประกอบด้วยบัตรรูปภาพที่จะต้องจัดลำดับเฉพาะให้เป็นเรื่องเป็นราว

5. *Object Assembly* เป็นการจัดให้เป็นรูปร่างของหุ่นวัตถุที่สมบูรณ์ภายในเวลาจำกัด การให้คะแนนสำหรับการปฏิบัติ คือ Arithmetic Digit Symbol, Block Design, Picture Arrangement และ Object Assembly ก็ดูที่ความรวดเร็วและความถูกต้องเป็นสำคัญ



เกณฑ์ปกติและการให้คะแนน

คะแนนดิบของ WAIS เปลี่ยนเป็นคะแนนมาตรฐาน โดยให้ mean = 10 และค่าความเบี่ยงเบนมาตรฐาน = 3 คะแนนจากสเกลนี้ได้มาจากกลุ่มขนาด 500 คน ซึ่งรวมบุคคลทุกประเภทที่มีอายุระหว่าง 20-34 ปี ส่วนกลุ่มตัวอย่างที่เอามาทำ norm นั้น ได้มาจากประชากร 1,700 คน มีชายหญิงเท่ากัน มีระดับอายุ 16-64 ปี

ค่าความเชื่อมั่น

แบบทดสอบเกือบทุก subtest ยกเว้น Digit Span และ Digit Symbol ที่หาโดยประมาณความสัมพันธ์ระหว่างคะแนนที่ได้จาก Digits Forward และ Digits Backward โดยทำจาก parallel form ส่วนแบบทดสอบอื่นๆ ใช้วิธี Split - half

ค่าของความเชื่อมั่นของ Full Scale ทุกๆ ระยะห่าง 3 ปี จะมีค่า .97 Verbal I.Q. จะมีค่าเท่ากับ .96 และ Performance I.Q. มีค่าเท่ากับ .93 และ .94 เหตุที่มีค่าสูงก็เนื่องจากมีค่าความมั่นคงภายใน (internal consistency) ในคู่มือของ WAIS ได้บอกค่าความคลาดเคลื่อนมาตรฐาน (Standard error of measurement) ว่า Verbal I.Q. มีค่า = 3 Performance I.Q. มีค่า = 4 ส่วน Full Scale I.Q. มีค่า = 2.6

4.3 Wechsler Pre - School and Primary Scale of Intelligence (WPPSI)

เป็นแบบทดสอบที่ใช้สำหรับเด็กเล็กๆ จัดพิมพ์ขึ้นในปี 1967 ใช้สำหรับเด็กอายุ 4-6½ ปี มีทั้งหมด 11 subtests และ 10 ฉบับใช้ในการหาค่าเชาวน์ปัญญา 8 ฉบับ ก็เลียนแบบมาจากของ wisc ส่วนอีก 3 ฉบับสร้างขึ้นมาใหม่ เพื่อใช้แทนแบบทดสอบของ wisc ที่ถามในด้านการหาเหตุผล

WPPSI แบ่งออกเป็น 2 ประเภทเช่นกัน คือ

- Verbal Scale
- Performance Scale

Verbal Scale ประกอบด้วย

1. ความรู้ทั่วไป
2. ศัพท์
3. เลขคณิต
4. ความคล้ายคลึงกัน
5. ความเข้าใจ

6. ถ้ามประเมิน (ใช้ถ้ามแทนแบบทดสอบ WISC ที่ถ้าม Digit Span) ก็คือ ผู้ดำเนินการสอบจะอ่านประโยคให้เด็กฟัง แล้วให้อ่านตามทันที แบบทดสอบนี้อาจให้เลือกทำหรืออาจให้ทำต่อจากแบบทดสอบที่ 5 เพื่อทำข้อมูลเพิ่มเติมให้มากยิ่งขึ้น แต่ไม่เอามารวมกับคะแนนรวมทั้งหมด

Performance Scale ประกอบด้วย

1. Animal House เป็นแผ่นป้ายที่มีรูปสัตว์ต่างๆ เช่น หมา ไก่ ปลา แมว ซึ่งสัตว์เหล่านั้นมีสีต่างๆ กัน เด็กก็ต้องเอาแท่งสีต่างๆ ใส่ลงไปในช่องที่รูปสัตว์เหล่านั้น โดยให้แท่งสีนั้นเป็นสีเดียวกันกับสัตว์แต่ละครั้ง การให้คะแนนก็จดเวลา สีที่ใส่ผิด

2. Picture Completion

3. Mazes

4. Geometric Design ให้เขียนรูปทรงเรขาคณิตตามแบบที่กำหนดให้ 10 รูป โดยให้ดินสอสี

แบบทดสอบ WPPSI นี้ ได้นำไปทดสอบกับเด็ก 1,200 คน โดยสอบจากเด็กตั้งแต่ช่วงอายุ 4-6½ ปี ในแต่ละช่วง ½ ปี ก็สอบเด็กชาย 100 คน หญิง 100 คน กลุ่มตัวอย่างสุ่มหาแบบแบ่งตามสภาพภูมิศาสตร์ ในเมือง-ชนบท คนขาว-ดำ อาชีพของบิดา คะแนนที่ได้มาทำเป็นคะแนนมาตรฐาน มี mean = 10 standard deviation = 3 ในแต่ละกลุ่มที่ห่างกัน ¼ ปี

คะแนนที่ได้ มีคะแนนรวมของ Verbal Scale คะแนนรวมของ Performance Scale และคะแนนรวมทั้งหมด (Full Scale) เมื่อมาทำเป็นคะแนนมาตรฐานแล้วมี mean = 100 s.d. = 15

การหาค่าความเชื่อมั่นของแบบทดสอบทุกฉบับ (ยกเว้น Animal House) หาโดยวิธี Split - half โดยใช้สูตรของ Spearman - Brown ส่วนคะแนนของแบบทดสอบ Animal House ใช้คิดจากเวลาที่เด็กใช้หาค่าความเชื่อมั่นโดยวิธีสอบซ้ำ เมื่อจบการสอบไปแล้วครั้งหนึ่ง

ค่าความเชื่อมั่นของคะแนนรวมทั้งหมด = .92 .93

ค่าความเชื่อมั่นของคะแนน Verbal = .87 .90

ค่าความเชื่อมั่นของคะแนน Performance = .84 .91

ค่าสหสัมพันธ์ภายในของแบบทดสอบ 11 ฉบับ มีค่าอยู่ระหว่าง .40 .60

5. TONI (Test of Nonverbal Intelligence) (A Language Free Measure of Cognitive Ability) (1982)

แบบทดสอบวัดเชาวน์ปัญญาฉบับนี้ สร้างขึ้นเมื่อปี 1982 โดยมีจุดประสงค์ที่จะหลักเลียงปัญหาเกี่ยวกับภาษาและวัฒนธรรม ผู้สร้างคือ Linda Brown, Rita J.Sherbenon และ Susan J.Dollar ลักษณะของ TONI จะเป็นภาพวัดความสามารถต่าง ๆ กัน คือ

1. Simple Matching
2. Analogy ซึ่งแยกเป็น
 - Addition
 - Subtraction
 - Progression
3. Classification
4. Interrection
5. Progression

วิธีการสร้าง ผู้สร้างจะศึกษาจากแบบทดสอบวัดสติปัญญาของ Raven Progressive Matrices, Performances Subtests ของ Wechsler's Scale และ Leiter International Performance Scale แล้วตรวจสอบกับผลงานของ Guilford (1956 a) นอกจากนั้น ก็ดูจากผลการทดลองของ Gagné (1959), Bourne (1963, 1967 ; Bourne & Gury 1968) Gberckberg 1(964) และ Dominowski 1(966) หลังจากนั้นจึงสร้างแบบทดสอบออกมา 307 ข้อ โดยให้มีเนื้อหาเป็นภาพ abstract ซึ่งไม่สื่อความหมายใด ๆ เลย เมื่อให้ผู้เชี่ยวชาญหลายๆ อาชีพตรวจสอบและวิจารณ์แล้ว ก็ได้ข้อสอบทั้งหมด 183 ข้อ นำไปทดสอบหาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ปรากฏว่ามีค่า $p = .15 \rightarrow .85$ และค่า r ตั้งแต่ $.30$ ถึง $.80$ เมื่อนำไปทดลองกับกลุ่มคน 3 กลุ่ม คือ กลุ่มเด็กอายุ 8 ขวบขึ้นไป ที่อยู่ในชั้น 1, 3, 5, 7, 9 กลุ่มหนุ่มสาว อายุ 18-35 ปี และกลุ่มคนแก่ อายุ 65-85 ปีขึ้นไป ปรากฏว่าได้ค่า $p = .15 \rightarrow .85$ และค่า $r .30 \rightarrow .80$ จากนั้นจึงแบ่งแบบทดสอบเป็น 2 ชุด คือ Form A 50 ข้อ และ Form B 50 ข้อ

การสร้างมาตรฐานของแบบทดสอบชุดนี้ ได้นำไปทดลองกับกลุ่มตัวอย่างในอเมริกา 28 รัฐ ได้กลุ่มตัวอย่างอายุระหว่าง 5 ปี ถึง 85 ปี 11 เดือน รวมทั้งสิ้น 1,929 คน

ค่าความเชื่อมั่นของแบบทดสอบนี้ เมื่อหาโดยวิธี α - Coefficient มีค่าระหว่าง $.78$ $.96$ หาโดยใช้สูตร KR_{21} ได้ค่าความเชื่อมั่น $.6 \rightarrow .9$ และเมื่อหาโดยใช้วิธีถ่วงน้ำหนักได้ค่าความเชื่อมั่น $.78 \rightarrow .95$

การหาค่าความเที่ยงตรง เมื่อใช้กับ Raven Progressive Matric (1983) กับคนหูหนวก ได้ค่าความเที่ยงตรงระหว่าง $.89 \rightarrow .92$ เมื่อใช้กับ WISC ในกลุ่มเด็กที่เรียนไม่ได้ และเด็กที่มีปัญหา มีค่าระหว่าง $.63 \rightarrow .92$ นอกจากนี้ ยังหาค่าความเที่ยงตรงกับแบบทดสอบมาตรฐานอื่น ๆ อีกหลายชนิด ก็จะมีค่าอยู่ระหว่าง $.63 \rightarrow .89$

การตีค่าความหมายของคะแนน แบ่งได้ดังนี้

Superior	คะแนน	131-145
สูงกว่าเฉลี่ย	”	116-130
ปานกลาง	”	85-115
ต่ำกว่าเฉลี่ย	”	70-84
Poor	”	55-69

จากค่าคะแนนดังกล่าว สามารถแปลงคะแนนเป็น QTONI, Z - score Stanine และ Percentile Rank ได้

อย่างไรก็ดี แบบทดสอบนี้ยังไม่แพร่หลายมากนัก แต่จากคุณสมบัติและวิธีดำเนินการสร้าง วิธีการสอบไม่ยุ่งยากเหมือนของ Stanford หรือของ Wechsler จึงคาดว่าพอจะนำมาใช้ประมาณค่าสติปัญญาของคนไทยได้ เนื่องจากแบบทดสอบนี้จัดปัญหาในเรื่องภาษา และวัฒนธรรมออกไปแล้ว

6. แบบทดสอบความถนัด (Aptitude Test)

ที่นำมาใช้กันทั่วไป ได้แก่

1. **SRA Primary Mental Abilities** เป็นแบบทดสอบชนิดหนึ่งที่ใช้กันทั่วไป เช่น The Chicago Tests of Primary Mental Abilities (PMA) ซึ่งพิมพ์ครั้งแรกในปี 1941 แบบทดสอบชุดนี้วัดตามทฤษฎีของ Thurstone ในแต่ละ factor Thurstone ได้เลือกเอาแบบทดสอบที่มีค่าความเที่ยงตรงสูงสุด แบบทดสอบนี้นำมาใช้กับโรงเรียนมัธยมและระดับวิทยาลัย ซึ่งแบ่งเป็น 6 ส่วน ในปี 1962 ก็ได้มีการปรับปรุงแบบทดสอบ PMA โดยแก้ไขข้อบกพร่อง แล้วแบ่งเป็น 5 ส่วน ใช้สอบกับนักเรียนตั้งแต่ชั้นอนุบาลจนถึงเกรด 12 คะแนนต่าง ๆ แบ่งเป็น 5 กลุ่ม ได้แก่ Verbal Meaning, Number Facility Reasoning Perceptual Speed และ Spatial Relation สำหรับ 5 factor ใช้ตั้งแต่เกรด 4-6 โดยยกเว้น Reasoning ไม่ต้องทดสอบในระดับต่ำกว่านี้ และ Perceptual Speed ก็ไม่ต้องสอบในชั้นสูงกว่าเกรด 6 คะแนนในแต่ละ factor นำมาแปลงให้เป็น Deviation I.Q.'s คู่มือแบบทดสอบก็จะอธิบายให้ครูผู้ดำเนินการสอบดำเนินการอย่างถูกต้อง สำหรับเด็กอนุบาลถึงเกรด 1 คะแนน ที่ได้จะให้เป็น ratio I.Q.'s เด็กเกรด 2-4 ก็ใช้ทั้ง ratio I.Q.'s และ Deviation I.Q.'s ส่วนคะแนนเด็กตั้งแต่เกรด 4 ขึ้นไป ก็ใช้เปอร์เซ็นต์ไทล์และ Stanine

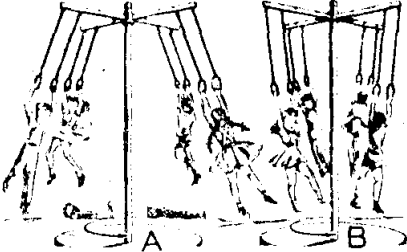
กลุ่มตัวอย่างของ PMA (1962) นั้นเป็นนักเรียน 32,393 คน จาก 93 โรงเรียน วิธีการเลือกกลุ่มตัวอย่างก็ทำแบบแบ่งเป็นชั้น ๆ ตามสภาพภูมิศาสตร์ และขนาดของโรงเรียน การหาค่าความเชื่อมั่นใช้วิธี Test-retest กับเด็กที่มีคุณลักษณะคล้ายคลึงกัน กลุ่มเล็ก ๆ ให้สอบซ้ำ

1 อาทิตย์ และ 4 อาทิตย์ ซึ่งมีค่าความเชื่อมั่นอยู่ระหว่าง .83-.95

2. **The Differential Aptitude Test (DAT)** พิมพ์ในปี 1947 แล้วนำมาปรับปรุงในปี 1963 แบบทดสอบชุดนี้สร้างขึ้นโดยมีจุดมุ่งหมายสำคัญที่จะใช้ในทางการศึกษา และการแนะแนว อาชีพ ให้เด็กเกรด 8-12 แบบทดสอบ DAT แบ่งคะแนนออกเป็น 8 กลุ่ม คือ Verbal Reasoning, Numerical Ability, Verbal Reasoning, Abstract Reasoning, Clerical Speed and Accuracy, Space Relations, Language Usage I (Spelling) และ Language Usage II (Grammar)

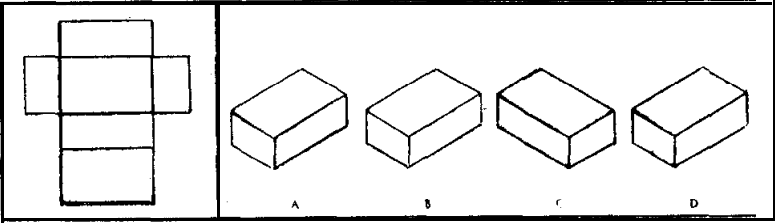
ตัวอย่าง

MECHANICAL REASONING
In which picture are the children whirling foster? (If equal, mark C.)



The correct answer is A.

SPACE RELATIONS
Which one of the following figures could be made by folding the pattern at the left? The pattern always shows the outside of the figure.
Note the grey surfaces.



The correct answer is D.

ต่อมาแบบทดสอบ DAT ได้มีการปรับปรุงขึ้นอีกในปี 1966 ได้มีการแก้ไขกระบวนการในการสร้างข้อสอบ Norm ก็ได้มาจากนักเรียน 50,000 คน จาก 195 โรงเรียนใน 43 รัฐ ค่าความเชื่อมั่นที่ได้ก็มีค่าสูงขึ้น การตีความหมายคะแนนใช้ทั้ง Percentile และ Stanine

3. Flanagan Aptitude Classification Tests (FACT) เป็นแบบทดสอบชุดหนึ่งที่วัดความถนัดใน factor ต่าง ๆ เริ่มสร้างหลังจากได้มีการทำวิจัยของ Flanagan เพื่อปรับปรุงกองทัพอากาศในสมัยสงครามโลกครั้งที่ 2 ส่วนใหญ่ของแบบทดสอบนี้เป็นการทดสอบเกี่ยวกับการให้คำปรึกษาทางด้านอาชีพ โดยให้ผู้ที่จะเข้าไปเป็นลูกจ้างเป็นผู้เลือก อาชีพต่าง ๆ ได้ถูกนำมาทำการ job analysis และได้องค์ประกอบของงาน 21 องค์ประกอบ ตัวอย่างเช่น

Assembly : ความสามารถที่จะเห็นขนาดรูปร่างของวัตถุนั้น โดยแยกจากวัตถุดิบอื่น ๆ ได้

Planning : ความสามารถในการวางแผนงาน การจัดการและการจัดตาราง เป็นความสามารถที่จะมองเห็นปัญหาที่อาจจะเกิดขึ้นเบื้องหน้าได้ และมีการคาดการณ์ล่วงหน้าว่า จะมีวิธีแก้ปัญหานั้นเป็นขั้นตอนอย่างไร

Ingenuity : ความสามารถในการสร้างสรรค์หรือมีทักษะในการประดิษฐ์ โดยแสดงความสามารถให้เห็นถึงวิธีการประดิษฐ์

แบบทดสอบชุดนี้มี 21 ชุด เป็นการทดสอบแบบ paper pencil test 19 ชุด ส่วนอีก 2 ชุด เป็นการทดสอบโดยให้ป้อน Tapping Test ซึ่งเป็น performance test แบบทดสอบ 19 ฉบับนั้นใช้เวลาทั้งหมด 10½ ชั่วโมง โดยอาจแบ่งเป็น 3 ตอน Norm ที่ได้เป็น national norm จากนักเรียน 11,000 คน ตั้งแต่เกรด 9 ถึง 12 ค่าความเชื่อมั่นของแบบทดสอบความถนัดทางด้านอาชีพมีค่าสูง ค่าความเที่ยงตรงก็หาโดยการติดตามผลงานทางด้านอาชีพที่เด็กเลือก ก็มีค่าสูง แบบทดสอบนี้จึงแสดงให้เห็นว่ามีค่า predictive validity ต่ออาชีพได้ดีมาก

4. Guilford - Zimmerman Aptitude Survey มีอยู่ 7 แบบทดสอบย่อย คือ

- Verbal Comprehension
- General Reasoning
- Numerical Operations
- Perceptual Speed
- Spatial Orientation
- Spatial Visualization
- Mechanical Knowledge

กิจกรรม

1. ให้นักศึกษาศึกษาลักษณะของแบบทดสอบวัดสติปัญญา แล้วเปรียบเทียบกับแบบทดสอบวัดความถนัด ว่ามีความแตกต่างกันอย่างไรบ้าง
2. ให้สำรวจรายชื่อแบบทดสอบวัดเชาวน์ปัญญาและแบบทดสอบวัดความถนัดที่พบอยู่ในประเทศไทย

แหล่งศึกษาค้นคว้าเพิ่มเติม

วิจิตธานี เจริญขวัญ การทดสอบทางจิตวิทยา (PC 303) กรุงเทพมหานคร สำนักพิมพ์มหา-
วิทยาลัยรามคำแหง 2532 หน้า 63-118

แบบฝึกหัดท้ายบท

1. จงอธิบายความเป็นมาของประวัติโครงสร้างทฤษฎีทางจิตวิทยา
2. Intelligence และ Aptitude มีความแตกต่างกันอย่างไรจงอธิบาย
3. แบบทดสอบ Stanford-Binet Intelligence Test มีลักษณะอย่างไร จงอธิบาย
4. แบบทดสอบวัดเชาวน์ปัญญาของ Wechsler มีกี่ประเภท อะไรบ้าง จงอธิบายความแตกต่างของแต่ละประเภท ตลอดจนวิธีการใช้แบบทดสอบนั้น ๆ
5. แบบทดสอบ TOM เหมาะที่จะใช้กับบุคคลประเภทใด มีลักษณะและแนวทางในการแปลความหมายผลการสอบอย่างไร
6. แบบทดสอบความถนัดที่นิยมใช้กันแพร่หลาย ได้แก่แบบทดสอบอะไรบ้าง จงอธิบาย