

151.2
๑488ก
ร.๓

ความสามารถในการตอบแบบทดสอบชนิดที่เป็นรูปภาพและที่เป็นตัวหนังสือ
ของนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7

THE LIBRARY
COLLEGE OF EDUCATION
BANGKOK, THAILAND

ปริญญานิพนธ์



เสนอต่อวิทยาลัยวิชาการศึกษา
เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต

13 สิงหาคม 2511

คณะกรรมการที่ปรึกษาประจำตัวนิติได้พิจารณาปัญหานี้จนบัดนี้แล้ว เห็นสมควร
รับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต ของวิทยาลัยวิชาการ
ศึกษาได้

๒๕๖๑ ๑๑/๑๒/๒๕๖๑ ประธาน

๐๙/๑๒/๒๕๖๑ กรรมการ

๑๕/๑๒/๒๕๖๑ กรรมการ

13 สิงหาคม 2511

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลงได้ เพราะได้รับความช่วยเหลืออย่างดียิ่งจาก
ผู้ช่วยศาสตราจารย์ ดร. ชวาล แพร่ตูล หัวหน้าสำนักงานทดสอบ วิทยาลัยวิชาการศึกษา
ประสานมิตร รองศาสตราจารย์ กุหลาบ มัดลิกะมาส และอาจารย์อารี สุทธิพันธุ์ ผู้เขียน
ขอกราบขอบพระคุณไว้อย่างสูง ณ ทนทวย

นอกจากนี้ ผู้เขียนขอขอบพระคุณ อาจารย์รัตนา ศิริพานิช อาจารย์ฉวน สายยศ
อาจารย์ณัฐชาติ สุวรรณวงศ์ ที่ช่วยตรวจแก้ไขคำภาษาไทยให้ดีขึ้น และอาจารย์พิตร ทองชั้น
ที่ได้อำนวยความสะดวกแบบทดสอบประเภทรูปภาพให้ดีขึ้น

ผู้เขียนขอขอบพระคุณ ท่านอาจารย์ใหญ่ ในโรงเรียนที่ได้ทำการศึกษาค้นคว้า ที่ได้
ให้ความร่วมมืออย่างดียิ่งในการ เก็บรวบรวมข้อมูลครั้งนี้ด้วย

ปริญญานิพนธ์ฉบับนี้สำเร็จเป็นรูปเล่มมาได้ เพราะได้รับความช่วยเหลืออย่างดียิ่ง
จากคุณ นิเวศน์ ชรรณรักษ์ ผู้พิมพ์ปริญญานิพนธ์ฉบับนี้ตลอดเล่ม ผู้เขียนขอขอบพระคุณไว้ ณ
ทนทวย

อังคณา ตันศิริพนานนท์

สารบัญ

บท	หน้า
1. บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายในการคนควา	4
ความสำคัญของการคนควา	5
ขอบเขตของการคนควา	6
คำนิยามศัพท์เฉพาะ	7
ประโยชน์จากการคนควา	7
2. เอกสารที่เกี่ยวข้องกับการศึกษานี้	9
3. วิธีดำเนินการ เก็บรวบรวมข้อมูลและ เทคนิคการคนควา	16
กลุ่มตัวอย่าง	16
เครื่องมือในการ เก็บรวบรวมข้อมูล	19
โดยใช้แบบทดสอบ	19
แบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ	19
แบบทดสอบไม่เข้าพวกประเภทรูปภาพ	20
แบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ	20
แบบทดสอบอุปมาอุปไมยประเภทรูปภาพ	21
การทดลองแบบทดสอบ	21
การดำเนินการสอบ	22
สถิติที่ใช้ในการวิเคราะห์ข้อมูลจากการศึกษานี้	23
4. ผลของการวิเคราะห์	27
สัญลักษณ์ที่ใช้ในการวิเคราะห์	28

ก. คุณภาพของแบบทดสอบในด้านการวัดความยากง่าย	29
ก. 1 คาสถิติพื้นฐานของแบบทดสอบ	29
ก. 2 การทดสอบความแตกต่างของคะแนนเฉลี่ย	31
ก. 3 คุณภาพของแบบทดสอบจำแนกเป็นรายชั้นเรียน	35
ก. 4 การทดสอบความแตกต่างของคะแนนเฉลี่ยจำแนกตามระดับชั้น	39
ก. 5 คุณภาพของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือจำแนกเป็นรายอายุ	42
ก. 6 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ จำแนกตามระดับอายุ	45
ก. 7 คุณภาพของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพจำแนกเป็นรายอายุ	47
ก. 8 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกตามระดับอายุ	49
ก. 9 คุณภาพของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือจำแนกเป็นรายอายุ	50
ก.10 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกตามระดับอายุ	52
ก.11 คุณภาพของแบบทดสอบไม่เข้าพวกประเภทรูปภาพจำแนกตามรายอายุ	53
ก.12 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกตามระดับอายุ	55
ก.13 คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 2 จำแนกตามเพศ	56

ก.14	คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 4	
	จำแนกตามเพศ	58
ก.15	คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 7	
	จำแนกตามเพศ	60
ข.	คุณภาพของแบบทดสอบในค่านอำนาจจำแนก	62
ข. 1	อำนาจจำแนกของแบบทดสอบโดยชั้นเรียน	64
ข. 2	อำนาจจำแนกของแบบทดสอบโดยอายุ	65
ข. 3	อำนาจจำแนกของแบบทดสอบโดยเพศ	66
ค.	สมการพยากรณ์ตามระดับชั้นเรียนและอายุ	67
ค. 1	สมการสำหรับพยากรณ์ระดับชั้นเรียน	67
ค. 2	สมการสำหรับพยากรณ์ตามระดับอายุ	71
5.	สรุปผลการวิเคราะห์ และข้อเสนอแนะ	75
	ความมุ่งหมายของการวิจัย	75
	กลุ่มตัวอย่างที่ใช้ในการศึกษา	75
	เครื่องมือในการเก็บข้อมูล	76
	การวิเคราะห์ข้อมูล	77
	ผลจากการวิเคราะห์ข้อมูล	77
	ข้อเสนอแนะ	80
	บรรณานุกรม	81
	ภาคผนวก	84
	ภาคผนวก ก	85
	ภาคผนวก ข	89

บัญชีตาราง

ตาราง	หน้า
1. จำนวนนักเรียนชั้นประถมศึกษาจากโรงเรียนที่ทำการศึกษา 12 แห่ง	16
2. จำนวนนักเรียนที่แยกตามระดับชั้นเรียนและเพศ	17
3. จำนวนนักเรียนที่แยกตามระดับอายุ	18
4. แสดงค่าความยากง่ายเฉลี่ย ความยากมาตรฐาน และค่าความเชื่อมั่น ของแบบทดสอบทั้ง 4 ฉบับ	22
5. ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ	30
6. ผลการวิเคราะห์ความแปรปรวนของแบบทดสอบ 2 ชนิดทั้ง 2 ประเภท	31
7. ผลต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ และระดับนัยสำคัญ ของผลต่างเหล่านั้น	33
8. ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ จำแนกเป็นรายชั้นเรียน	36
9. ผลการวิเคราะห์ความแปรปรวนของคะแนนสอบทั้ง 4 ฉบับ จำแนกตามระดับชั้น ประถมศึกษาปีที่ 2, ประถมปีที่ 4, และประถมศึกษาปีที่ 7	39
10. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ ในแต่ละชั้นเรียน และระดับความเชื่อมั่นของผลต่างเหล่านั้น	41
11. ค่าสถิติพื้นฐานของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ จำแนกเป็น รายอายุ	43
12. ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบอุปมาอุปไมย ประเภทตัวหนังสือ จำแนกเป็นรายอายุ	45

13.	ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภท ตัวหนังสือในแต่ละอายุ และความเชื่อมั่นของผลทางเหล่านั้น	46
14.	ค่าสถิติพื้นฐานของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกเป็น รายอายุ	47
15.	ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบอุปมาอุปไมย ประเภทรูปภาพ จำแนกตามรายอายุ	49
16.	ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภท รูปภาพ ในแต่ละอายุและระดับความเชื่อมั่นของผลทางเหล่านั้น	50
17.	ค่าสถิติพื้นฐานของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ จำแนกเป็น รายอายุ	51
18.	ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบไม่เข้าพวก ประเภทตัวหนังสือ จำแนกตามรายอายุ	52
19.	ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกเป็นตัวหนังสือ ในแต่ละอายุและความเชื่อมั่นของผลทางเหล่านั้น	53
20.	ค่าสถิติพื้นฐานของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกเป็นรายอายุ	54
21.	ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบไม่เข้าพวกประเภท รูปภาพ จำแนกตามรายอายุ	55
22.	ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ ในแต่ละอายุและความเชื่อมั่นของผลทางเหล่านั้น	56
23.	ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ ของนักเรียนชั้นประถมศึกษาปีที่ 2 จำแนกตามเพศ	57

24.	คำสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ ของนักเรียนชั้นประถมศึกษาปีที่ 4 จำแนกตามเพศ	59
25.	คำสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ ของนักเรียนชั้นประถมศึกษาปีที่ 7 จำแนกตามเพศ	61
26.	ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามระดับชั้นเรียน	64
27.	ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามระดับอายุ	65
28.	ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามเพศ	66
29.	การหา R_p และสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับตามระดับชั้นเรียน	69
30.	การหา R_p และสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับ ตามระดับอายุ	72
31.	ค่า P_H , P_L , p , r และ Δ ที่ได้จากการวิเคราะห์แบบทดสอบ อุปมาอุปไมย	90
32.	ค่า P_H , P_L , p , r และ Δ ที่ได้จากการวิเคราะห์แบบทดสอบไม่ ไม่เข้าพวก	91
33.	การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกิติ ของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ	92
34.	การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกิติ ของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ	93
35.	การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกิติ ของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ	94

36.	การแจกแจงความถี่, เปอร์เซนต์ และคะแนน T - ประติ ของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ	95
37.	ค่า $N, \sum X, \sum X^2$ ของแบบทดสอบทั้ง 4 ฉบับ จำแนกตามระดับชั้นเรียน	96
38.	ค่า $N, \sum X, \sum X^2$ ของแบบทดสอบทั้ง 4 ฉบับ จำแนกตามรายอายุ	96
39.	ค่า $N, \sum X, \sum X^2$ ของแบบทดสอบทั้ง 4 ฉบับ แยกตามความแตกต่าง ระหว่างเพศ	97
40.	ค่า $\sum XY$ ของแบบทดสอบทั้ง 4 ฉบับ	97
41.	ค่า intercorrelation ของแบบทดสอบทั้ง 4 ฉบับ	98
42.	ค่า $\sum X$ ในการวิเคราะห์ความแปรปรวนระหว่างชนิดและประเภทของ แบบทดสอบควย Analysis of variance แบบ 2 x 2 Factorial design	98
43.	ค่า $\sum X$ ในการวิเคราะห์ความแปรปรวนระหว่างแบบทดสอบกับระดับ ชั้นเรียน แบบ 4 x 3 Factorial design	99
44.	ค่า t ในการทดสอบนัยสำคัญของคะแนนเฉลี่ยระหว่างอายุของ แบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ	99
45.	ค่า t ในการทดสอบนัยสำคัญของคะแนนเฉลี่ยระหว่างอายุของ แบบทดสอบอุปมาอุปไมยประเภทรูปภาพ	100
46.	ค่า t ในการทดสอบนัยสำคัญของคะแนนเฉลี่ยระหว่างอายุของ แบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ	101

47.	ค่า t ในการทดสอบนัยสำคัญของคะแนนเฉลี่ยระหว่างอายุของ แบบทดสอบไม่เข้าพวกประเภทรูปภาพ	102
48.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ ตามรายชั้นเรียน	102
49.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ ตามรายชั้นเรียน	103
50.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ ตามรายชั้นเรียน	103
51.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ ตามรายชั้นเรียน	104
52.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมย ประเภทตัวหนังสือ ตามรายอายุ	105
53.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมย ประเภทรูปภาพ ตามรายอายุ	106
54.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ ตามรายอายุ	107
55.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ ตามรายอายุ	108

56.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ จำแนกตามเพศ	109
57.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกตามเพศ	109
58.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ จำแนกตามเพศ	110
59.	สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกตามเพศ	110



บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1. ภาพการเปรียบเทียบความสามารถของนักเรียนชั้นประถมศึกษาในแต่ละวิชา.....

34



คำนำ

ภูมิหลัง

เมื่อกล่าวถึงการทดสอบ บุคคลส่วนใหญ่จะนึกถึงการสอบที่ใช้แบบทดสอบประเภทภาษา (Verbal test) ซึ่งประกอบด้วยข้อคำถามที่ใช้อักษรและภาษาเป็นหลัก และผู้สอบก็จะต้องเขียนตอบบรรยายในกระดาษ ถึงแม้ในปัจจุบันนี้ทั้งการศึกษาและงานวิจัยในแบบทดสอบที่ให้ผู้สอบเลือกหาคำตอบที่ถูกต้องที่สุดจากบรรดาคำตอบที่กำหนดให้ไว้ก่อนแล้ว (Fixed-response test) ก็ตามทั้งตัวคำถามและตัวเลือกเหล่านั้น ก็ยังคงใช้อักษรและภาษาเป็นพื้นฐาน อย่างไรก็ตาม แบบทดสอบเหล่านี้คงจัดว่า เป็นแบบทดสอบประเภทภาษาทั้งสิ้น จริงอยู่แม้ว่าในบางครั้ง เราใช้ภาษาเป็นเครื่องมือสื่อความหมายแสดงความคิดที่ติดต่อกันมากมาย และแรงจูงใจ¹ ได้ทำการค้นคว้าทางภาษาพบว่า มนุษย์ใช้ภาษาในบางครั้ง ๆ ด้านการฟัง การพูด การอ่าน และการเขียน มีจำนวนเปอร์เซ็นต์ถึง 42, 32, 15 และ 11 ตามลำดับ แต่ผลที่แรงจูงใจพบนั้นไม่ได้เป็นเครื่องยืนยันว่า การทดสอบที่ใช้ภาษาจะไต่ตามความมุ่งหมายครบ 100 % ทั้งนี้เพราะว่าภาษาที่ใช้ในแบบทดสอบนั้นอาจจะสูงกวาระดับของผู้เรียนซึ่งทำให้ผู้เรียนไม่เข้าใจภาษาที่ใช้ในแบบทดสอบนั้น.

ในทศวรรษที่แล้วมานี้ ปรากฏว่ามีผู้นิยมสร้างแบบทดสอบขึ้นมาอีกประเภทหนึ่ง เป็นแบบทดสอบที่ทั้งข้อคำถามและคำตอบประกอบด้วยรูปภาพ (Nonverbal test) เป็นส่วนใหญ่ ความมุ่งหมายของแบบทดสอบประเภทนี้มุ่งที่จะให้เกิดความยุติธรรมแก่ผู้เข้าสอบที่มีทักษะทางด้านภาษาแตกต่างกัน ในกรณีของการทดสอบนั้นไม่ต้องการให้ข้อผิดพลาดทางภาษาเข้ามาเกี่ยวข้อง ดังเช่นในการสอบวัดสมรรถภาพด้านมิติสัมพันธ์และด้านเหตุผล เป็นต้น.

¹ Rankin, Paul T., "Listening Ability," Proceeding of The Ohio State Educational Conference Ohio State University, pp. 172 - 183, 1929.

การพัฒนาการของแบบทดสอบจากประเภทภาษาเป็นประเภทรูปภาพนั้น เป็นแนวโน้มซึ่ง
ให้เห็นความก้าวหน้าของวิทยาการด้านการวัดผลการศึกษาเป็นอย่างดี ขวาล แพทย์กุล²
เคยเน้นถึงวิธีการสร้างคำถามโดยใช้รูปภาพว่า "เพราะรูปภาพสามารถช่วยให้เข้าใจคำถามง่าย
และชัดเจนตรงกันมากขึ้น. ฉะนั้น, จึงสามารถทำให้ข้อสอบมีประสิทธิภาพสูงขึ้นเองโดยปริยาย."

การศึกษาในประเทศไทย โดยเฉพาะในระดับประถมศึกษา ปรากฏว่า นักเรียนใน
ระดับนี้สอบตกกันมากทีเดียว เป็นผลให้เกิดการสูญเสียเวลาทางการศึกษา ซึ่งยอมรับกระเทือน
ต่อเศรษฐกิจของประเทศ ในรายงานของกองการประถมศึกษา กรมสามัญศึกษาประจำปี 2508³
แสดงว่า สาเหตุที่นักเรียนในระดับประถมศึกษา โดยเฉพาะอย่างยิ่ง ชั้นประถมศึกษาปีที่ 1 ถึง
ชั้นประถมศึกษาปีที่ 4 สอบตกกันมาก ก็คือ นักเรียนมีความรู้ภาษาไทยไม่เพียงพอ เฉลี่ยแล้ว
นักเรียนเหล่านี้มีความรู้ภาษาไทยไม่ถึงระดับตามต้องการถึงร้อยละ 16.74

การที่นักเรียนสอบตกและต้องเรียนซ้ำชั้นนั้น อาจมีสาเหตุเนื่องมาจาก นักเรียนไม่เข้าใจ
ภาษาที่ใช้ในแบบทดสอบ หรือแบบทดสอบใช้ภาษาเข้าใจยากไม่เหมาะกับระดับของผู้เรียน หรือภาษา
ที่ใช้มีความหมายหลายแง่หลายมุมไม่รัดกุมเพียงพอ ทำให้เกิดปัญหาแก่ผู้ถูกทดสอบ⁴ และยังไม่
สามารถวัดให้ครอบคลุมความรู้ของผู้เรียนในระดับนั้น ๆ. เมื่อเป็นเช่นนั้น แบบทดสอบที่ใช้แตกภาษา
อย่างเดียวกัน อาจจะขาดความเที่ยงตรง และไม่สามารถวัดความสำเร็จทางการศึกษาได้ จึงทำให้
นักเรียนสอบตกกันมากถึงครึ่งโหลมาแล้ว.

ในต่างประเทศได้มีการวิจัยเกี่ยวกับการพัฒนาทางภาษาในระดับชั้นประถมศึกษากันมาก
แอนเดรีย⁵ ให้ความเห็นว่า การใช้ภาษาเป็นเรื่องของการเรียนรู้ ซึ่งต้องอาศัยทั้งวุฒิภาวะ
และประสบการณ์ของแต่ละคน นอกจากนี้ เปียเจต์⁶ ยังได้ศึกษาความเข้าใจทางภาษา และให้

² ขวาล แพทย์กุล, เทคนิคการวัดผล พ.ศ. 2510 หน้า 166

³ กองการประถมศึกษา, กรมสามัญศึกษา รายงานประจำปี 2508 หน้า 60 - 64

⁴ ขวาล แพทย์กุล, ล.ค. หน้า 154 - 155

⁵ Andreas, Berton G ; Experimental Psychology p. 595.

⁶ Piaget, Jean, The Language and Thought of The Child, p. 133.

คำอธิบายโดยสรุปว่า พัฒนาการทางภาษาของคนจะเริ่มจากส่วนรวมไปหาส่วนย่อย. เด็กทารกสามารถจะออกเสียงต่าง ๆ ได้มากมายโดยไม่รู้ความหมาย. ภายหลังเมื่ออายุเพิ่มขึ้นมีพัฒนาความรู้ทางภาษาสูงขึ้น ก็จะรู้จักแยกแยะการเรียนรู้ความหมายของคำต่าง ๆ เพื่อสนองความต้องการของตน ส่วนคำที่ไม่มีความหมายก็จะลดน้อยลง. ดังนั้นสำหรับเด็กชั้นประถมศึกษา จึงควรอย่างยิ่งที่จะส่งเสริมพัฒนาการทางการเรียนรู้โดยใช้รูปภาพ และการทดสอบก็ควรจะใช้แบบทดสอบที่เป็นรูปภาพบ้าง เพราะรูปภาพมีความหมายอยู่ในตัว นักเรียนเข้าใจง่าย แม้ว่าจะยังไม่มีความรู้ทางภาษาอย่างดีพอก็ตาม อีกประการหนึ่งการเรียนรู้ไม่จำเป็นจะต้องใช้ภาษาหรือตัวหนังสือเสมอไป อาจใช้รูปภาพโยงไปหาตัวหนังสือก็ได้ ดังที่ คีส⁷ ทำการศึกษาโดยใช้วิธีการเรียนรู้แบบรูปภาพโยงไปหาตัวหนังสือที่เรียกว่า วิธีการเรียนรู้แบบคู่สัมพันธ์ (Paired associate learning) ผลปรากฏว่าได้ผลดีกว่าวิธีการเรียนรู้โดยใช้ภาษาแต่อย่างเดียว.

จากการศึกษาเกี่ยวกับแบบทดสอบชนิดต่าง ๆ นักการศึกษาหลายท่านมีความเห็นตรงกันว่าแบบทดสอบที่เป็นตัวหนังสือเมื่อนำมาใช้กับเด็กที่เรียนช้า และระดับอายุต่ำเกินไปมักจะไม่มีใครได้ผล นักเรียนจะสับสนตกกันมาก⁸ จึงสู้แบบทดสอบที่เป็นรูปภาพไม่ได้ ดังที่แอสแมนและกล๊อค⁹ เคยทำการศึกษามาแล้ว และสรุปว่าแบบทดสอบรูปภาพเหมาะสมสำหรับเด็กที่เรียนช้าและระดับอายุต่ำ.

⁷ Deese, Jean , The Psychology of Learning, pp. 166 - 167.

⁸ Anastasi, Anne , Psychological Testing, p. 239.

⁹ Ahmann, J. Stanley and Glock, Marvin D., Evaluation Pupil Growth, p. 69.

ในหลักสูตรชั้นประถมศึกษาตอนต้น พ.ศ. 2503¹⁰ มีความมุ่งหมายในการสอนภาษาไทยมีอยู่ข้อหนึ่งว่า "การสอนภาษาไทยในห้องเรียนเพื่อส่งเสริมพัฒนาทางภาษาของผู้เรียนให้เกิดความรู้ ทักษะ และเจตคติที่เหมาะสมกับวัยให้สามารถใช้ภาษาไทยได้ดี ทั้งการฟัง การพูด การอ่าน และการเขียน เพื่อจะใช้เป็นเครื่องมือในการติดต่อและแสวงหาความรู้" ดังที่ไจ้กล่าวในตอนต้นแล้วว่า เด็กในระดับประถมศึกษาตอนต้น การพัฒนาทางภาษายังไม่เจริญเพียงพอ ยังจะต้องเรียนรู้อีกมากทั้งภายในห้องเรียนและนอกห้องเรียน ดังนั้นการที่ทดสอบเด็กในระดับนี้ ด้วยแบบทดสอบตัวหนังสือ จึงไม่สามารถที่จะวัดได้เต็มที่ตามความมุ่งหมายในหลักสูตรที่กล่าวไว้ จึงควรวัดผลด้วยแบบทดสอบที่เป็นรูปภาพให้เหมาะสมกับระดับชั้นเรียน วัย และความแตกต่างระหว่างเพศด้วย เพื่อที่จะได้ผลถูกต้องตามสภาพความเป็นจริง แต่อย่างไรก็ตามแบบทดสอบที่เป็นตัวหนังสือก็ดี และแบบทดสอบที่เป็นรูปภาพก็ดีต่างก็เป็นประโยชน์ต่อขบวนการศึกษา และช่วยในการประเมินผลการศึกษาได้ถูกต้อง

ตามหลักฐานและเหตุผลต่าง ๆ ดังไจ้กล่าวมาแล้วข้างต้น จึงเห็นสมควรจะมีการทดสอบด้วยแบบทดสอบทั้งสองประเภทนี้ควบคู่กันไปทุกครั้งที่มีการวัดผลในระดับประถมศึกษา จึงจะสามารถวัดผลการศึกษาในระดับนี้ได้เหมาะสมและรัดกุมยิ่งขึ้น

ความมุ่งหมายในการตรวจ

๑ ในการตรวจครั้งนี้อยู่ศึกษาตามหัวข้อต่อไปนี้

1. เพื่อให้ทราบถึงความสามารถในการตอบแบบทดสอบที่เป็นรูปภาพของนักเรียนชั้นประถมศึกษาชั้นนี้มีความเชื่อมั่นและเที่ยงตรงปานใด เมื่อเปรียบเทียบกับความสามารถในการตอบแบบทดสอบที่เป็นตัวหนังสือ และผลที่ได้จากการทดสอบนักเรียนทั้งสองแบบนี้ประเภทไหนมีความเชื่อมั่นและเที่ยงตรงกว่ากัน

2. เพื่อให้ทราบว่า เด็กระดับชั้นเรียนใด ควรทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือ

3. เพื่อให้ทราบว่า เด็กระดับอายุเท่าไร ควรทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือ

4. เพื่อให้ทราบว่า ความแตกต่างระหว่างเพศว่า จะมีผลต่อการทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือหรือไม่ และเพื่อเปรียบเทียบว่าเพศใดสามารถดีกว่ากันในการทำแบบทดสอบแต่ละประเภท

5. เพื่อหาสมการพยากรณ์ระดับชั้นเรียนและระดับอายุต่าง ๆ ของนักเรียนชั้นประถมศึกษาที่ใช้แบบทดสอบนี้

ความสำคัญของการค้นคว้า

1. การศึกษาครั้งนี้ทำให้ทราบว่า แบบทดสอบที่เป็นรูปภาพหรือที่เป็นตัวหนังสือใช้กับนักเรียนชั้นประถมศึกษาใดผลเที่ยงตรง และเชื่อมั่นได้มากที่สุด

2. ทำให้ทราบว่า เด็กระดับชั้นเรียนใด ควรทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือ จึงจะเหมาะสมที่สุด

3. ทำให้ทราบว่า เด็กอายุเท่าไร ควรทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือจึงจะได้ผลเที่ยงตรงที่สุด

4. ทำให้ทราบว่า เพศใดที่สามารถทดสอบด้วยแบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือได้ดี

5. ได้ตัวพยากรณ์ชุดหนึ่ง อันจะสามารถใช้เป็นเครื่องมือในการพยากรณ์ว่า นักเรียนคนใดอยู่ในระดับชั้นเรียนใดหรืออายุเท่าไร

6. การวิจัยครั้งนี้ จะเป็นส่วนหนึ่งของข้อเท็จจริงที่แสดงให้เห็นถึงจุดเด่นจุดด้อยของการสร้างแบบทดสอบทั่วไป และเป็นแนวทางสำหรับปรับปรุงการสร้างแบบทดสอบให้เกิดผลดีต่อการศึกษา

ขอบเขตของการค้นคว้า

การศึกษาครั้งนี้อยู่ในขอบเขตจำกัดดังนี้ คือ

1. การศึกษาครั้งนี้จะศึกษาเฉพาะนักเรียนประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7 การที่เลือกระดับชั้นเรียนทั้ง 3 ชั้นนี้ เพราะว่าชั้นประถมศึกษาปีที่ 2 เป็นชั้นที่เริ่มจะอ่านหนังสือได้แล้ว ส่วนชั้นประถมศึกษาปีที่ 4 และปีที่ 7 นี้ เป็นชั้นตัวประโยคประถมศึกษา จึงเลือกระดับชั้นเรียนเหล่านี้เป็นกลุ่มตัวอย่างที่จะทำการศึกษา.
2. จำนวนนักเรียนที่ทำการศึกษามีทั้งหมด 1316 คน แยกตามระดับชั้นเรียนคือ ประถมศึกษาปีที่ 2 มี 424 คน ประถมศึกษาปีที่ 4 มี 467 คน และประถมศึกษาปีที่ 7 มี 425 คน.
3. กลุ่มตัวอย่างเลือกมาจากโรงเรียนชั้นประถมศึกษาที่สังกัดเทศบาลนครกรุงเทพ 3 โรงเรียน โรงเรียนสังกัดกรมสามัญศึกษา 3 โรงเรียน โรงเรียนประชาสาคิต 2 โรงเรียน และโรงเรียนราษฎร์ 4 โรงเรียน รวมเป็น 12 โรงเรียน.
4. แบบทดสอบที่จะใช้กับนักเรียนในกลุ่มตัวอย่างที่เลือกมานี้ เป็นแบบทดสอบความถนัดทางการเรียนประเภทรูปภาพ และประเภทตัวหนังสือชนิดจัดเข้าพวกและอนุภาคูปไมย ชนิดละ 2 ฉบับคือ
 แบบทดสอบประเภทรูปภาพชนิดจัดเข้าพวก
 แบบทดสอบประเภทรูปภาพชนิดอนุภาคูปไมย
 แบบทดสอบประเภทตัวหนังสือชนิดจัดเข้าพวก
 แบบทดสอบประเภทตัวหนังสือชนิดอนุภาคูปไมย
 การที่ใช้แบบทดสอบชนิดจัดเข้าพวก และชนิดอนุภาคูปไมยนั้น ก็ควยเหตุว่าเป็นแบบทดสอบที่วัดคานเหตุผลโดยตรง ซึ่งจะทำให้การวิจัยในครั้งนี้เที่ยงตรงและถูกต้องยิ่งขึ้น
5. นักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7 เป็นนักเรียนที่ได้ทำการศึกษามาแล้วถึงภาคเรียนที่สาม ของปีการศึกษา 2510.

คำนิยามศัพท์เฉพาะ

1. แบบทดสอบที่เป็นรูปภาพ (Nonverbal test) หมายถึงแบบทดสอบที่เป็นรูปภาพที่มีความหมาย ซึ่งสามารถแปลความหมายและตีความหมายได้ตรงกัน.
2. แบบทดสอบที่เป็นตัวหนังสือ (Verbal test) หมายถึงแบบทดสอบที่ใช้ภาษาเป็นสื่อแสดงความคิดระหว่างมนุษย์.
3. ความเที่ยงตรง (Validity) หมายถึงคุณสมบัติที่จะทำให้ผู้ไขว่คว้าถึงวัตถุประสงค์หรือแบบทดสอบที่สามารถวัดในสิ่งที่ต้องการจะวัด ได้อย่างถูกต้องตามความมุ่งหมายที่กำหนดไว้.
4. ความเชื่อมั่น (Reliability) หมายถึงแบบทดสอบที่ได้คะแนนคงที่แน่นอนไม่แปรผันไปมา ถึงแม้จะทำการทดสอบซ้ำอีกในเด็กกลุ่มเดิม คะแนนที่ได้ของแต่ละคนก็ยังคงที่อยู่ หรือใกล้เคียงกับคะแนนเดิม.
5. กลุ่มตัวอย่าง (Sample) หมายถึงนักเรียนที่กำลังทำการศึกษาในครั้งนั้น.
6. ตัวพยากรณ์ (Predictor) หมายถึงเครื่องมือชุดหนึ่งที่ใช้พยากรณ์นักเรียนชั้นประถมศึกษาว่าอยู่ระดับชั้นเรียนใด หรืออายุเท่าไรอันประกอบด้วย แบบทดสอบที่เป็นรูปภาพชนิดจัดเข้าพวก แบบทดสอบที่เป็นรูปภาพชนิดอุปมาอุปไมย, แบบทดสอบที่เป็นตัวหนังสือชนิดจัดเข้าพวกและแบบทดสอบที่เป็นตัวหนังสือชนิดอุปมาอุปไมย.

ประโยชน์จากการค้นคว้า

ประโยชน์และคุณค่าโดยตรงต่อการศึกษาที่จะได้รับจากการศึกษานี้

1. เพื่อประโยชน์ในการสร้างแบบทดสอบว่าควร เป็นแบบที่เป็นรูปภาพ หรือที่เป็นตัวหนังสือในการทดสอบนักเรียนในชั้นประถมศึกษา
2. เพื่อประโยชน์ในการสร้างแบบทดสอบว่า ระดับชั้นเรียนใดในชั้นประถมศึกษาควรทดสอบด้วยแบบทดสอบที่เป็นรูปภาพ หรือแบบทดสอบที่เป็นตัวหนังสือ จึงจะวัดได้เที่ยงตรงตามระดับชั้นเรียนของนักเรียน

3. เพื่อประโยชน์ในการสร้างแบบทดสอบที่จะใช้ในการทดสอบนักเรียนชั้นประถมศึกษาว่า อายุเท่าไร ควรทดสอบควยแบบทดสอบที่เป็นรูปภาพ และอายุเท่าไร ควรทดสอบควยแบบทดสอบที่เป็นตัวหนังสือ.

4. เพื่อประโยชน์ในการสร้างแบบทดสอบสำหรับนักเรียนชั้นประถมศึกษาว่าควรใช้แบบทดสอบที่เป็นรูปภาพหรือแบบทดสอบที่เป็นตัวหนังสือให้เพียงตรงตามความแตกต่างระหว่างเพศ.

5. เพื่อประโยชน์ในการพยากรณ์ระดับชั้นเรียนและระดับอายุของนักเรียนในการทำแบบทดสอบ.

6. เพื่อเป็นข้อ เสนอแนะในการปรับปรุงการวัดผลนักเรียนให้แกครูในโรงเรียนประถมศึกษา

7. ผลจากการศึกษาครั้งนี้ อาจใช้เป็นข้อ เสนอแนะในการ เขียนแบบเรียนและเอกสาร วิชาการอื่น ๆ สำหรับนักเรียนชั้นประถมศึกษา ว่า ควรมีรูปภาพประกอบให้มาก หรือเป็นตัวหนังสือ ล้วน ๆ ตามความเหมาะสม.

เอกสารที่เกี่ยวข้องกับการค้นคว้า

นักการศึกษาในสหรัฐอเมริกาให้ความสนใจเกี่ยวกับการวัดผลการศึกษานักเรียน ด้วยวิธีการวัดต่าง ๆ มาก โดยเฉพาะการวัดผลที่ใช้แบบทดสอบหลายประเภท เช่น ประเภทที่เป็นตัวหนังสือและประเภทที่เป็นรูปภาพ นักการศึกษาหลายคนได้ทำการวิจัยเปรียบเทียบเพื่อหาความเหมาะสมในการทดสอบนักเรียน พร้อมทั้งอธิบายความหมายของแบบทดสอบทั้ง 2 ชนิด ดังที่แอสแมน และกลอค¹ ได้ให้ความหมายของแบบทดสอบที่เป็นตัวหนังสือว่า เป็นแบบทดสอบผลสัมฤทธิ์ที่ใช้กระดาษ - ดินสอ (paper and pencil achievement test) อาศัยการเขียนและภาษาพูดในการตอบปัญหาแต่ละข้อของข้อสอบ ส่วนกรีนและคณะ² ได้ให้ความหมายของแบบทดสอบที่เป็นรูปภาพว่า เป็นแบบทดสอบกระดาษ - ดินสอเช่นเดียวกับแบบทดสอบที่เป็นตัวหนังสือ แตกต่างกันตรงที่แบบทดสอบที่เป็นรูปภาพไม่ต้องใช้ภาษา และการเขียนในการตอบปัญหาแต่ละข้อ เพราะรูปภาพมีความหมายอยู่ในตัวแล้ว นักเรียนสามารถที่จะตอบทันทีโดยไม่ต้องอาศัยความเข้าใจในภาษา.

ในปี ค.ศ. 1958 เกนเดอร์เซนและเฟลด์ท³ ต่างก็มีความสนใจ และศึกษาการหาสหสัมพันธ์ระหว่างแบบทดสอบที่เป็นตัวหนังสือ แบบทดสอบรูปภาพกับแบบทดสอบผลสัมฤทธิ์หลังจากที่ไบเจอร์, สเตราค, บลูมเมอร์ และเลสเบอร์ ได้กล่าวไว้ว่า แบบทดสอบที่เป็นตัวหนังสือมีสหสัมพันธ์กับแบบทดสอบผลสัมฤทธิ์มากกว่าแบบทดสอบที่เป็นรูปภาพ แต่แม็คลีนและชาเบอร์ เลท

¹ Ahmann, J. Stanley and Glock, Marvin L., Evaluating Pupil Growth, p. 69.

² Greene, Harry A., Lorgensen, Albert N., and Gerberich, J. Raymon, Measurement and Evaluation in The Elementary School, pp. 41 - 42.

³ Gendersen, Richard, D., and Feldt, Leonard S., "The Relationship of Difference Between Verbal and Nonverbal Intelligence Score to Achievement," Journal of Educational Psychology 51 : 115 - 121, 1960.

ได้ศึกษาค้นคว้า แบบทดสอบที่เป็นรูปภาพมีสหสัมพันธ์กับแบบทดสอบผลสัมฤทธิ์สูงกว่าแบบทดสอบตัวหนังสือ ทั้งเกณฑ์คะแนนและเฟลตทต่างก็เริ่มหาข้อเท็จจริง จึงได้ร่วมกันทำการศึกษาโดยใช้กลุ่มตัวอย่าง จากโรงเรียนชั้นประถมศึกษาปีที่ 4 จำนวน 522 คน เครื่องมือที่ใช้ในการศึกษาค้นคว้านี้ใช้แบบทดสอบ ของแคลิฟอร์เนีย พบว่ากลุ่มที่ทำแบบทดสอบตัวหนังสือได้ดี จะสามารถทำแบบทดสอบที่เป็นคำศัพท์ และภาษาได้ดี ส่วนกลุ่มที่ทำแบบทดสอบรูปภาพได้ดี จะสามารถทำแบบทดสอบวิชาคณิตศาสตร์ได้ดี.

เฮจและสตรูด⁴ ได้ทำการศึกษาเช่นกันพบว่า การอ่านจะมีสหสัมพันธ์กับแบบทดสอบ ตัวหนังสือสูงกว่าแบบทดสอบที่เป็นรูปภาพ.

คอมมูกโนว⁵ ได้ศึกษาหาความสัมพันธ์ระหว่างแบบทดสอบที่เป็นตัวหนังสือกับแบบทดสอบ ผลสัมฤทธิ์ในการเรียนของนักเรียนมีค่า .68 - .94 ซึ่งมีค่าสหสัมพันธ์เฉลี่ยเป็น .86 ส่วนสหสัมพันธ์ ระหว่างแบบทดสอบที่เป็นรูปภาพกับผลสัมฤทธิ์ทางการเรียน .67 - .92 และมีสหสัมพันธ์เฉลี่ยเป็น .89 จากการศึกษาเหล่านี้แสดงว่า แบบทดสอบที่เป็นรูปภาพก็ใช้ในการวัดผลนักเรียนได้ดี และ ไม่จำเป็นจะต้องใช้แบบทดสอบที่เป็นตัวหนังสือเสมอไป.

ในการศึกษาต่อมาได้มีการศึกษาเพื่อหาว่า แบบทดสอบประเภทใดที่ใช้ได้เหมาะสมกับ นักเรียนทุกสังคมและทุกชาติได้ ซึ่งจากการศึกษาของนักวิจัยหลายท่านพบว่า แบบทดสอบที่เป็นรูปภาพ จะใช้ได้ในทุกสังคมและทุกชาติได้ ส่วนแบบทดสอบที่เป็นตัวหนังสือจะใช้ได้กับชาติใดชาติหนึ่งหรือ สังคมใดสังคมหนึ่งโดยเฉพาะ เพราะภาษาหรือตัวหนังสือของแต่ละชาติและสังคมไม่เหมือนกัน. นอกจากนี้แบบทดสอบที่เป็นรูปภาพยังสามารถที่จะใช้ทดสอบเปรียบเทียบระหว่างกลุ่มที่มีวัฒนธรรม ต่างกัน ภาษาต่างกันให้สอดคล้องกับสภาพความเป็นจริงของผู้สอบ⁶ ตั้งในปี ค.ศ. 1917 สหรัฐอเมริกาได้ใช้แบบทดสอบที่เป็นรูปภาพที่เรียกว่า อาร์มมี เบตา (Army Beta)

⁴ Hage, Deen, S; and Stroud, James B; "Reading Proficiency and Intelligence Scores, Verbal and Nonverbal," Journal of Educational Research 52 : 258 - 262, March, 1959.

⁵ Green, Edward B., Measurements of Behavior, pp. 89 - 90.

⁶ Anastasi, Anno, Psychological Testing, pp. 240 - 241.

เพื่อคัดเลือกทหารจากบุคคลหลายชาติหลายภาษาสำหรับที่จะวัดว่า บุคคลใดมีสติปัญญาสูงพอที่จะเป็นทหารที่มีไหวพริบดีในการสงครามได้⁷

เอกสารวิจัยต่างประเทศหลายฉบับได้กล่าวถึงการศึกษาเรื่องการใช้แบบทดสอบที่เป็นตัวหนังสือและรูปภาพกับนักเรียนที่มีหลายภาษา และอยู่ในสังคมเท่า ๆ กัน เช่น โบเกอร์⁸ ศึกษาจากกลุ่มนักเรียนชั้นประถมศึกษาปีที่ 1, 2, 3 ในเมืองเวอร์จิเนีย มีเด็กผิวขาวจำนวน 53 คน ผิวดำจำนวน 51 คน และใช้แบบทดสอบสติปัญญาประเภทตัวหนังสือของโอทิส (Otis Quick - Scoring Mental Ability Test) กับแบบทดสอบสติปัญญาประเภทรูปภาพของแคลิฟอร์เนีย (California Mental Maturity Test) ปรากฏผลดังนี้,

1. นักเรียนที่มาจากสังคมในชนบทได้คะแนนต่ำกว่าคะแนนเฉลี่ยของแบบทดสอบสติปัญญาทั้ง 2 ชุด.
 2. กลุ่มผิวดำได้คะแนนแบบทดสอบที่เป็นรูปภาพสูงกว่าคะแนนแบบทดสอบตัวหนังสืออย่างมีนัยสำคัญ
 3. กลุ่มผิวขาวได้คะแนนแบบทดสอบตัวหนังสือดีพอ ๆ กับคะแนนของแบบทดสอบที่เป็นรูปภาพ
- นอกจากนี้ คิลเลล⁹ ได้ทำการศึกษาการใช้แบบทดสอบที่เป็นตัวหนังสือกับที่เป็นรูปภาพในกลุ่มตัวอย่างที่มีสองภาษาในเชื้อชาติต่างกัน คือใช้กลุ่มตัวอย่างในโรงเรียนประถมศึกษาปีที่ 3 ของเมืองเบอร์กลีย์ จำนวน 41 คน และในเมืองแคลิฟอร์เนีย จำนวนอีก 42 คน โดยแยกเป็นกลุ่มตัวอย่างที่ใช้สองภาษามีจำนวน 42 คน เป็นเด็กชาย 24 คน หญิง 18 คน และอีก 41 คน

②-7 Ross, C.C., Measurement in Today's School, pp. 44 - 46.

⁸ Boger, Jack H., "An Experimental Study of The Effect of Perceptual Training On Group I.Q. Test Score of Elementary Pupils in Rural Ungraded School," Journal of Educational Research 46 : 43 - 52, September, 1952.

①-9 Killell, Jack E., "Bilingualism And Language Nonlanguage Intelligence Score of Third - Grade Children," Journal of Educational Research 52 : 263 - 268, March, 1959.

ใช้ภาษาเดียว (Unilingual) เป็นเด็กชาย 30 คน หญิง 11 คน ซึ่งกลุ่มตัวอย่างเหล่านี้ประกอบด้วยเชื้อชาติต่าง ๆ ดังนี้,

จีน 19.0 %

ญี่ปุ่น 16.7 %

ฝรั่งเศส 9.5 %

เยอรมัน 4.8 %

อิตาลี 4.8 %

ชาวคิซ 4.8 %

ชาติอื่น ๆ นอกเหนือจากที่กล่าวมานี้ 2.4 %

เครื่องมือที่ใช้ในการวิจัยคือ แบบทดสอบแคลิฟอร์เนีย ฉบับ 1953 (California Short Form Test of Mental Maturity) กับแบบทดสอบการอ่านของแคลิฟอร์เนีย (California Reading Test Primary Form A.A.) ได้ผลดังนี้คือ

1. เด็กชั้นประถมศึกษาปีที่สามที่ใช้สองภาษาในชีวิตประจำวัน ได้คะแนนทางด้านภาษาต่ำกว่ากลุ่มที่ใช้ภาษาเดียว
2. เด็กชั้นประถมศึกษาปีที่สามที่ใช้ภาษาเดียวได้คะแนนทางภาษาสูงกว่าได้คะแนนจากแบบทดสอบที่เป็นรูปภาพ
3. กลุ่มที่ใช้ภาษาเดียวที่พ่อแม่มีฐานะปานกลาง ได้คะแนนทางภาษาสูงกว่ากลุ่มที่ใช้สองภาษาที่พ่อแม่มีฐานะปานกลางเช่นกัน
4. กลุ่มที่ใช้สองภาษา และกลุ่มที่ใช้ภาษาเดียวต่างมีความแตกต่างในด้านความสามารถในการอ่านอย่างไม่มีนัยสำคัญ
5. กลุ่มที่ใช้สองภาษา ได้คะแนนของแบบทดสอบรูปภาพสูงกว่าแบบทดสอบทางภาษา การศึกษาข้างต้นนี้แสดงให้เห็นว่า แบบทดสอบที่เป็นรูปภาพใช้ได้ดีกว่าแบบทดสอบที่เป็นภาษา หรือตัวหนังสือสำหรับเด็กกลุ่มที่อยู่ในสังคมต่างกัน และที่มีการใช้ภาษาตลอดจนเชื้อชาติต่างกัน

ในการเปรียบเทียบความแตกต่างระหว่างเพศในการใช้แบบทดสอบที่เป็นตัวหนังสือนั้น เอ็ดมันด์¹⁰ ได้ทำการศึกษาในเรื่องนี้ โดยการใช้กลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่สิบเอ็ด จากโรงเรียนมัธยมศึกษา 66 โรงเรียน ในจำนวนนี้มี 63 โรงเรียน ที่สุ่มมาจากโรงเรียนในชนบท และทำการศึกษาโดยใช้แบบทดสอบสติปัญญาของโอทิส (Otis Mental Ability Test) ปรากฏผลดังนี้,

1. ในชั้นที่อยู่ในสังคมเศรษฐกิจต่ำที่สุด จะมีความแตกต่างในแบบทดสอบสติปัญญาด้านภาษา ระหว่างชายและหญิงอย่างมีนัยสำคัญ.
2. เด็กชายที่อยู่ในสังคมเศรษฐกิจสูงสุด ได้คะแนนมากกว่าเด็กชายที่อยู่ในสังคมเศรษฐกิจต่ำสุด ซึ่งแตกต่างอย่างมีนัยสำคัญ.
3. เด็กชายที่มีสถานะทางเศรษฐกิจสูงมักจะทำคะแนนได้มากกว่าเด็กหญิงที่มีสถานะเศรษฐกิจปานกลาง และยังทำคะแนนได้มากกว่าเด็กหญิงที่มีสถานะเศรษฐกิจสูงเล็กน้อยอีกด้วย.

ส่วนการศึกษาทางด้านอายุก็ได้มีนักวิจัยหลายคนได้ทำการศึกษาเช่นกัน ไดแก เจอร์ซิลด์¹¹ ทำการศึกษาเกี่ยวกับการใช้แบบทดสอบที่เป็นตัวหนังสือและพบว่า เด็กจะทำแบบทดสอบประเภทนี้ได้ดีตามระดับอายุ นอกจากนี้ ท่านผู้นี้ยังได้อ้างถึงการศึกษาของสมิทซึ่งเป็นการศึกษาว่าอเมริกันว่า พัฒนาการทางภาษาของเด็กจะเพิ่มขึ้นตามอายุ ยิ่งมีอายุมากก็จะมีความสามารถทางภาษามาก ส่วนอะนาสตาสี¹² กล่าวว่าเด็กที่มีอายุน้อยหรือเด็กที่มีการเรียนรู้ทางภาษาไม่สามารถทำแบบทดสอบที่เป็นตัวหนังสือได้ดีเท่าที่ควร เพราะยังมีความเข้าใจทางภาษาน้อย เด็กเหล่านี้เหมาะที่จะทดสอบด้วยแบบทดสอบที่เป็นรูปภาพ เพราะแบบทดสอบประเภทนี้ไม่ต้องการความสามารถทางภาษา ก็สามารถที่จะทำได้

¹⁰ Edmonds, William S., "Sex Difference in The Verbal Ability Socio - Economically Depressed Group," The Journal of Educational Research, 58 : 61 - 64, October, 1964.

¹¹ Jersild, Arthur T., Child Psychology, p. 403.

¹² Anastasi, Anne op. cit. p. 239.

ใน ค.ศ. 1932 ภูเก็ต 13 ได้ทำการศึกษาเพื่อที่จะทราบวาระวัยอายุใดเหมาะสำหรับแบบทดสอบที่เป็นตัวหนังสือหรือที่เป็นรูปภาพ และได้สร้างแบบทดสอบมินนิโซตาชั้นฉบับหนึ่ง สำหรับทดสอบเด็กก่อนเข้าโรงเรียน (Minnesota Preschool Test) สำหรับเด็กที่มีอายุตั้งแต่ $1\frac{1}{2}$ - 6 ปี แบบทดสอบชุดนี้แบ่งออกเป็น 2 ฉบับ คือ ฉบับหนึ่งเป็นแบบทดสอบที่เป็นตัวหนังสือ อีกฉบับหนึ่งเป็นแบบทดสอบที่เป็นรูปภาพ ทั้ง 2 ฉบับนี้ต่างก็นำมาทดสอบกับเด็กที่มีอายุต่ำกว่า 3 ปี ปรากฏว่า ทั้งแบบทดสอบที่เป็นตัวหนังสือและที่เป็นรูปภาพต่างก็ใช้ไม่ได้ผลในระดับเด็กที่มีอายุต่ำกว่า 3 ปี แต่ใช้ได้ดีเฉพาะแบบทดสอบที่เป็นรูปภาพกับเด็กที่มีอายุมากกว่า 3 ปีขึ้นไป

ส่วน การที่ 14 ไม่เพียงแต่ทำการศึกษาในการใช้แบบทดสอบที่เป็นตัวหนังสือและที่เป็นรูปภาพกับระยะวัยเท่านั้น ยังได้ศึกษาว่า แบบทดสอบประเภทใดเหมาะสำหรับเด็กชั้นใด โดยที่ใช้กลุ่มตัวอย่างจากนักเรียนชั้นประถมศึกษาปีที่ 5 และปีที่ 6 จำนวน 235 คน ซึ่งแยกเป็นนักเรียนชาย 117 คน และนักเรียนหญิง 118 คน ของโรงเรียนประถมศึกษาในเมืองนิวยอร์ก เครื่องมือที่ใช้ในการทดสอบเป็นแบบทดสอบทางภาษาของพินต์เนอร์ฉบับบี (The Pintner General Ability, Form B) และแบบทดสอบรูปภาพของพินต์เนอร์ฉบับเค (Form K) ปรากฏผลการศึกษาดังนี้

1. แบบทดสอบที่เป็นรูปภาพได้คะแนนเฉลี่ยสูงกว่าคะแนนเฉลี่ยของแบบทดสอบที่เป็นภาษาถึง 8.28 คะแนน
2. เด็กที่มีอายุ 126.54 เดือน จะได้คะแนนแบบทดสอบที่เป็นรูปภาพสูงกว่าเด็กอายุ 117.92 ถึง 8.62 คะแนน
3. เด็กในชั้นประถมศึกษาปีที่ 5 และปีที่ 6 ไม่มีความแตกต่างอย่างมีนัยสำคัญระหว่างแบบทดสอบที่เป็นตัวหนังสือและที่เป็นรูปภาพ
4. ค่าสหสัมพันธ์ระหว่างแบบทดสอบที่เป็นตัวหนังสือกับที่เป็นรูปภาพ มีค่า .58

13 Green, Edward B., op. cit. p. 88.

14 Darcy, Natalie T., "The Performance of Bilingual Puerto Rican Children On Verbal And On Nonlanguage Test of Intelligence," Journal of Educational Research 45 : 499 - 506, March, 1952.

รอส 15 ยังได้ศึกษาเกี่ยวกับระดับชั้นเรียนเหมือนกัน พบว่าควรวิเคราะห์แบบทดสอบที่เป็น
รูปภาพกับเด็กที่อยู่ในชั้นประถมศึกษาปีที่หนึ่ง และเด็กที่เริ่มเข้าโรงเรียน เพราะเด็กในระดับนี้
ยังอ่านหนังสือไม่ออกตลอดจนยังเขียนหนังสือไม่เป็น จึงควรวิเคราะห์แบบทดสอบรูปภาพว่าการทดสอบ
เด็กในระดับนี้

เอกสารการค้นคว้าของนักวิจัยเหล่านี้ เป็นส่วนหนึ่งที่ผู้เขียนได้นำมาเป็นแนวความคิดในการที่
จะศึกษาค้นคว้า เพื่อให้ได้มีประโยชน์ต่อกระบวนการศึกษาต่อไป.



วิธีการเก็บรวบรวมข้อมูลและเทคนิคการคำนวณ

กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการศึกษานี้ได้แก่นักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7 จำนวนทั้งหมด 1316 คน จากโรงเรียนต่าง ๆ 4 ประเภท คือโรงเรียนประถมศึกษา 2 โรงเรียนในสังกัดเทศบาลนครกรุงเทพ 3 โรงเรียนในสังกัดกรมสามัญศึกษา 3 โรงเรียนราษฎร์ 4 โรงเรียนรวมโรงเรียนที่ทำการสุ่มตัวอย่างทั้งหมด 12 โรงเรียนด้วยกัน โดยมีจำนวนนักเรียนชาย และนักเรียนหญิงตามโรงเรียนทั้ง 12 โรงเรียนดังในตารางที่ 1

ตารางที่ 1 จำนวนนักเรียนชั้นประถมศึกษาจากโรงเรียนที่ทำการศึกษา 12 แห่ง

โรงเรียน	สังกัด	ป 2		ป 4		ป 7		รวม	%
		ชาย	หญิง	ชาย	หญิง	ชาย	หญิง		
ประถมศึกษา ประสานมิตร	ฝึกหัดครู	29	31	20	14	26	16	136	10.33
ประถมศึกษา สวนสุนันทา	ฝึกหัดครู	12	25	14	30	7	37	125	9.50
วิชาสุทอง	เทศบาลฯ	17	15	14	23	22	16	107	8.13
วัดลาดบัวขาว	เทศบาลฯ	21	12	16	22	26	10	107	8.13
สวนลุมพินี	เทศบาลฯ	10	8	14	13	20	18	83	6.31
วีรานุวัตร	สามัญ	24	15	25	21	21	10	116	8.81
วัดชนะสงคราม	สามัญ	15	18	24	24	20	15	116	8.81
ทุ่งมหาเมฆ	สามัญ	10	20	16	17	22	14	99	7.52
ลาซาล	ร.ร.ราษฎร์	38	—	62	—	42	—	142	10.79
อุดมศาสตร์	ร.ร.ราษฎร์	17	15	19	10	6	8	75	5.70
พรวิทยา	ร.ร.ราษฎร์	23	19	16	23	21	17	119	9.04
ครูณิเวศน์	ร.ร.ราษฎร์	16	14	19	11	13	18	91	6.91
รวม		232	192	259	208	246	179	1316	100
%		17.63	14.59	19.68	15.81	18.69	13.60	100	

จากตารางที่ 1 จะเห็นว่า จำนวนนักเรียนชายจะมีมากกว่านักเรียนหญิง คือชั้นประถมศึกษาปีที่ 2 จะมีจำนวนนักเรียนชาย 232 คน คิดเป็น 17.63 % นักเรียนหญิง 192 คนคิดเป็น 14.59 % ชั้นประถมศึกษาปีที่ 4 มีจำนวนนักเรียนชาย 259 คน คิดเป็น 19.68 % นักเรียนหญิง 208 คน คิดเป็น 15.81 % ชั้นประถมศึกษาปีที่ 7 มีจำนวนนักเรียนชาย 246 คน คิดเป็น 18.69 % นักเรียนหญิง 179 คน คิดเป็น 13.60 %

ในการวิเคราะห์ข้อมูล ผู้เขียนจะแยกตามระดับชั้นเรียน และเพศตามตารางที่ 2

ตารางที่ 2 จำนวนนักเรียนที่แยกตามระดับชั้นเรียน และเพศ

ระดับชั้นเรียน	ชาย	หญิง	รวม	%
ป 2	232	192	424	32.22
ป 4	259	208	467	35.49
ป 7	246	179	425	32.29

นอกจากจะแยกตามระดับชั้นเรียนและเพศแล้ว ยังแยกตามระดับอายุ เพื่อสะดวกต่อการวิเคราะห์ข้อมูลตามตารางที่ 3

ตารางที่ 3 จำนวนนักเรียนที่แยกตามระดับอายุ

อายุ	ป 2		ป 4		ป 7		รวม	%
	ชาย	หญิง	ชาย	หญิง	ชาย	หญิง		
7 ปี	41	33	—	—	—	—	74	5.62
8 ปี	74	77	2	1	—	—	154	11.70
9 ปี	59	43	32	23	—	—	157	11.90
10 ปี	36	26	84	61	—	—	208	15.81
11 ปี	16	11	79	76	2	5	189	14.36
12 ปี	6	2	46	28	43	32	157	11.93
13 ปี	—	—	14	11	11	104	211	16.03
14 ปี	—	—	2	5	5	58	108	8.21
15 ปี	—	—	—	3	3	39	58	4.41
รวม	232	192	259	208	246	179	1316	100
%	17.63	14.59	19.68	19.81	18.69	13.60	100	

เครื่องมือในการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการศึกษาค้นคว้าครั้งนี้ใช้แบบทดสอบความถนัดทางการเรียน ซึ่งแบ่งเป็น 2 ประเภทคือ แบบทดสอบความถนัดประเภทตัวหนังสือ กับแบบทดสอบความถนัดประเภทรูปภาพ ทั้งแบบทดสอบความถนัดประเภทตัวหนังสือและประเภทรูปภาพ ต่างก็แบ่งเป็น 2 ชุดคือ แบบทดสอบชุดไม่เข้าพวก กับแบบทดสอบอุปมาอุปไมย รวมทั้งหมดมีแบบทดสอบ 4 ฉบับคือ แบบทดสอบตัวหนังสือชนิดไม่เข้าพวก แบบทดสอบตัวหนังสือชนิดอุปมาอุปไมย แบบทดสอบรูปภาพ ชนิดไม่เข้าพวกและแบบทดสอบรูปภาพชนิดอุปมาอุปไมย แบบทดสอบทั้งหมั้้นสร้างโดยอาศัยหลักของแบบทดสอบความถนัดทางการเรียนที่มาตรฐานแล้วเป็นตัวอย่าง แบบทดสอบความถนัดทางการเรียนที่ใช้เป็นตัวอย่างคือ แบบทดสอบฟลานาแกน แอบติจูด คลาสซิฟิเคชัน เทสต์ (Flanagan Aptitude Classification Test) แบบทดสอบเอส อาร์ เอ พิคเจอเรียล รีซันนิง เทสต์ (S.R.A. Pictorial Reasoning Test) แบบทดสอบแคลิฟอร์เนีย เทสต์ ออฟ เมนตัล แมทิวริตี (California Test of Mental Maturity) การสร้างแบบทดสอบทั้ง 4 ฉบับนี้ผู้เชี่ยวชาญทางการวัดผลการศึกษาและผู้เชี่ยวชาญทางด้านภาษาเป็นผู้แก้ไขปรับปรุงก่อน แล้วจึงนำไปทดลองกับกลุ่มตัวอย่าง แบบทดสอบในการใช้เก็บรวบรวมข้อมูล จึงมีลักษณะดังนี้.

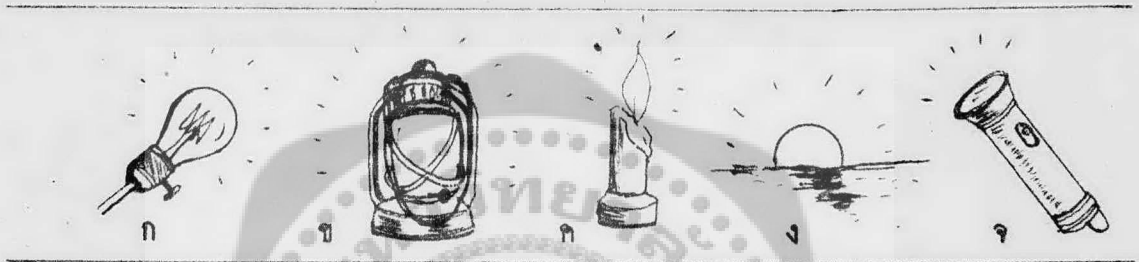
1. แบบทดสอบไม่เข้าพวกที่เป็นตัวหนังสือ เป็นแบบทดสอบปรนัยชนิดเลือกตอบ (Multiple choice) มี 40 ข้อ แต่ละข้อมีตัวเลือก (option) อยู่ 5 ตัว ใน 5 ตัวเลือกนี้ จะมีอยู่ 4 ตัวเลือกที่มีลักษณะหรือความหมายบางอย่างร่วมกัน คล้ายกัน หรือมีความสัมพันธ์เป็นประเภทเดียวกัน สกุดเดียวกัน และจะมีอยู่ข้อหนึ่งที่แตกต่างออกไปไม่เข้าพวกกับข้ออื่น ๆ ดังตัวอย่าง,

ข้อ (0) ข้อใดที่ไม่เข้าพวก ?

- ก. ไฟฉาย
- ข. ตะเกียง
- ค. เทียนไข
- ง. กวงอาทิตย์
- จ. หลอดไฟฟ้า

ข้อ ก. ง. หรือกวงอาทิตย์ไม่เข้าพวก เพราะเป็นสิ่งที่เกิดขึ้นเองตามธรรมชาติ ส่วนข้ออื่น ๆ เป็นสิ่งทมนุษย์ประดิษฐ์ขึ้น.

2. แบบทดสอบไม่เข้าพวกประเภทรูปภาพ เป็นแบบทดสอบปรนัยชนิดเลือกตอบมี 40 ข้อ และมีลักษณะ เหมือนกับแบบทดสอบตัวหนังสือชนิดไม่เข้าพวกทุกประการ เว้นแต่ว่า แบบทดสอบฉบับหนึ่ง เป็นตัวหนังสือ ส่วนอีกฉบับหนึ่งเป็นรูปภาพ คือแต่ละข้อจะมีรูปภาพอยู่ 5 ภาพ ใน 5 ภาพนี้ จะมี อยู่ 4 ภาพ ที่มีลักษณะหรือความหมายบางอย่างร่วมกัน คล้ายกัน หรือมีความสัมพันธ์เป็นประเภท เดียว สกปรกเดียวกัน และจะมีอยู่รูปภาพหนึ่งที่แตกต่างกันออกไป ดังตัวอย่าง



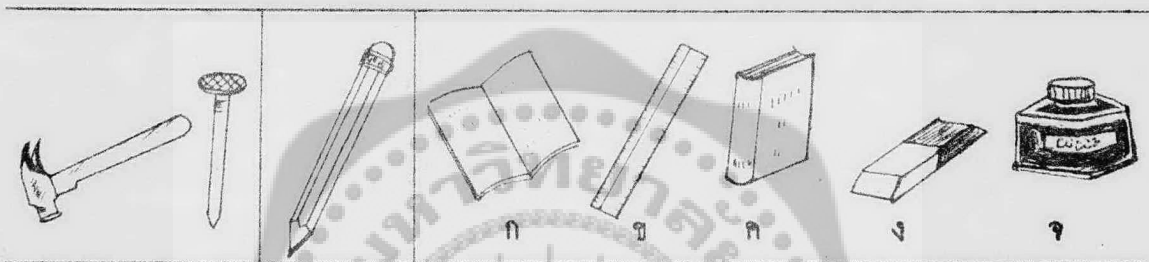
3. แบบทดสอบอุปมาอุปไมยที่เป็นตัวหนังสือ เป็นแบบทดสอบปรนัยชนิดเลือกตอบมีอยู่ 40 ข้อ เป็นแบบทดสอบที่มีอยู่ 2 ข้อที่มีความสัมพันธ์กันอย่างไรก็ตามหนึ่งทางซ้ายใหญ่ก่อน จากนั้นในข้อที่ 3 ว่าสัมพันธ์กับข้อใดใน 5 ข้อ ในลักษณะเดียวกับ 2 ข้อแรกทางซ้ายมือ ดังตัวอย่าง

(0) คอน : ตะปู —————> คินสอ : ?

- ก. สมุด
- ข. หมึก
- ค. ยางลบ
- ง. หนังสือ
- จ. ไม่บรรทัด

ข้อนี้ตอบ ข้อ ก. เพราะคินสอจะตรงกับสมุด ซึ่งมีความสัมพันธ์ทำนองเดียวกับคอน ตรงกับตะปู

4. แบบทดสอบรูปมาอุปไมยที่เป็นรูปภาพ เป็นแบบทดสอบปรนัยชนิดเลือกตอบที่เป็นรูปภาพ มีอยู่ 40 ข้อ แบบทดสอบฉบับนี้เหมือนกับแบบทดสอบตัวหนังสือชนิดรูปมาอุปไมยทุกประการ ผิดไปแต่ว่า เป็นแบบทดสอบรูปภาพ ซึ่งมีภาพ 2 ภาพทางซ้ายมือใหญ่ก่อน ภาพทั้ง 2 นี้จะมีบางสิ่งเกี่ยวข้องกับทางใจทางหนึ่ง จากนั้นให้ดูภาพที่ 3 ว่าสัมพันธ์เป็นทำนองเดียวกับ 2 ภาพแรกจาก 5 ภาพทางขวามือ ดังตัวอย่าง,



การทดลองแบบทดสอบ

สร้างแบบทดสอบแต่ละชนิด ๆ ละ 70 ข้อ แล้วนำไปทดลองกับกลุ่มตัวอย่างคือนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7 ของโรงเรียนวัดใหม่ทองถมจำนวน 174 คน กับโรงเรียนวัดโบสถ์จำนวน 93 คน รวมเป็น 267 คน แล้วนำแบบทดสอบมาวิเคราะห์เป็นชนิด ๆ ตามเทคนิคการวิเคราะห์ข้อทดสอบ 27 % ทั้งกลุ่มสูงและกลุ่มต่ำ แล้วเปิดตารางสำเร็จรูปของ จุง - เทห์ - แฟน¹ เพื่อหาความยากง่าย อำนาจจำแนกและความยากมาตรฐานเป็นรายข้อ (Δ) แล้วเลือกเฉพาะข้อสอบที่มีความยากง่ายตั้งแต่ .20 - .80 และอำนาจจำแนกตั้งแต่ .20 ขึ้นไปจำนวน 40 ข้อในแบบทดสอบแต่ละชนิด, แล้วหาความยากง่ายเฉลี่ย (p เฉลี่ย) ต่อจากนั้นหาค่าความเชื่อมั่นของข้อสอบด้วยวิธีของคูเคอร์² ซึ่งแสดงไว้ในตารางที่ 4,

¹ Fan, Chung - Teh, Item Analysis Table, pp. 6 - 32.

² Garrett, Henry E., Statistics in Psychology and Education, pp. 341 - 342.

ตาราง 4 แสดงค่าความยากง่ายเฉลี่ย ความยากมาตรฐาน และค่าความเชื่อมั่น
ของแบบทดสอบทั้ง 4 ฉบับ

ประเภทของแบบทดสอบ	ความยากง่ายเฉลี่ย	ความยากมาตรฐาน	ค่าความเชื่อมั่น
ตัวหนังสือชนิดไม่เข้าพวก	49.4500	13.0525	.83
ตัวหนังสือชนิดอุปมาอุปไมย	47.7750	13.2350	.86
รูปภาพชนิดไม่เข้าพวก	49.4500	13.0525	.74
รูปภาพชนิดอุปมาอุปไมย	47.7750	13.2350	.81

การดำเนินการสอบ

เมื่อเลือกข้อทดสอบใดชนิดละ 40 ข้อ และนำไปปรับปรุงให้ดีขึ้น ก็พิมพ์แบบทดสอบใหม่
ชนิดละ 500 ฉบับ รวม 2000 ฉบับ แล้วนำไปทดสอบกับกลุ่มตัวอย่างที่ได้เลือกแล้ว 12 โรงเรียน
ด้วยกันคือ

1. โรงเรียนประถมสาธิต ประสานมิตร
2. โรงเรียนประถมสาธิต สวนสุนันทา
3. โรงเรียนวัดธาตุทอง
4. โรงเรียนวัดลาดบัวขาว
5. โรงเรียนสวนกุหลาบ
6. โรงเรียนวิจิตรวาทธรรม
7. โรงเรียนวัดชนะสงคราม
8. โรงเรียนทุ่งมหาเมฆ
9. โรงเรียนลาซาล
10. โรงเรียนอุดมศึกษา

11. โรงเรียนพรวิทยา

12. โรงเรียนครุณิเวศน์

ทั้ง 12 โรงเรียนนี้ ผู้เขียนไปดำเนินการสอบด้วยตนเอง เพื่อป้องกันความผิดพลาดต่าง ๆ ที่จะเกิดขึ้นในระหว่างดำเนินการสอบ

เมื่อทดสอบได้ครบทุก 12 โรงเรียนแล้ว ผู้เขียนได้ตรวจกระดาษคำตอบเหล่านี้ด้วยตนเอง ในการตรวจกระดาษคำตอบได้ใช้แผนภูมิแจกพลาสติก (Key) เจาะรูตรงข้อเลือกที่ถูก แล้วทาบไปบนแผนกระดาษคำตอบ โดยให้ข้อละหนึ่งคะแนนเหมือนกันหมดทุกแบบทดสอบ คำตอบที่ผิดหรือไม่ได้ตอบ หรือชี้คำตอบข้อละเกินหนึ่งชี้ก็ถือว่าผิดไม่ให้คะแนน

สถิติที่ใช้ในการวิเคราะห์ข้อมูลจากการศึกษา

1. รายเฉลี่ยของคะแนน คะแนนที่ได้จากแบบทดสอบทั้ง 4 ชนิด นำมาหารรายเฉลี่ยที่จะได้นำไปเปรียบเทียบกันระหว่างแบบทดสอบ และระหว่างกลุ่มตามต้องการ การหารรายเฉลี่ยใช้สูตร

$$\bar{X} = \frac{\sum X}{N}$$

เมื่อ $\bar{X}$ แทนคะแนนเฉลี่ย
 $\sum X$ แทนผลรวมทั้งหมดของคะแนน
 N แทนจำนวนกลุ่มตัวอย่างในกลุ่มที่ทดลอง

2. ความเบี่ยงเบนมาตรฐาน (Standard deviation) ตัวเลขสถิตินี้ บอกปริมาณการกระจายของคะแนน ถ้ามีค่าสูงก็แปลว่า คะแนนที่นักเรียนสอบได้มีค่าแตกต่างกัน ถ้ามีค่าน้อยก็แปลว่า นักเรียนได้คะแนนใกล้เคียงกันมาก เด็กมีความเก่งอ่อนปาน ๆ กัน การคำนวณความเบี่ยงเบนมาตรฐานใช้สูตร

$$s = \sqrt{\frac{N \sum X^2 - (\sum X)^2}{N(N-1)}}$$

³ Guilford, J.P., Fundamental Statistics in Psychology and Education, p. 44.

⁴ Spiegel, Murray R., Theory and Problems of Statistics, p. 77.

$\sum s$	แทน ความเบี่ยงเบนมาตรฐาน
$\sum X$	แทน ผลรวมทั้งหมดของคะแนน
$\sum X^2$	แทน ผลรวมทั้งหมดของคะแนนแต่ละตัวยกกำลังสอง
N	แทน จำนวนนักเรียนในกลุ่มนั้น

3. สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) เป็นการหาความสัมพันธ์ของข้อมูล 2 ชนิดที่คนเดียวกันสอบ ถ้าสัมประสิทธิ์สหสัมพันธ์สูง แสดงว่า ของสองสิ่งขึ้นลงตามกัน
ใช้สูตร

$$r = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

$\sum r$	แทน สัมประสิทธิ์สหสัมพันธ์
$\sum XY$	แทน ผลรวมของผลคูณคะแนน X กับ Y
$\sum X \sum Y$	แทน ผลคูณของผลรวมคะแนน X กับผลรวมคะแนน Y
$\sum X, \sum Y$	แทน ผลรวมทั้งหมดของคะแนน X, คะแนน Y
$\sum X^2, \sum Y^2$	แทน ผลรวมทั้งหมดของคะแนน X, Y แต่ละตัวยกกำลังสอง
N	แทน จำนวนนักเรียนทั้งหมดของกลุ่มนั้น

4. วิเคราะห์ความแปรปรวน (Analysis of Variance) เพื่อดูว่า ส่วนรวมของคะแนนของตัวแปรทั้ง 14 ชนิด ว่าแตกต่างกันอย่างมีนัยสำคัญหรือไม่.

⁵ Garrette , Henry E., op. cit. p. 143.

5. ทดสอบความแตกต่างโดยวิธี t - test เพื่อหาว่าคะแนนเฉลี่ยของตัวแปรแต่ละชนิดนั้นแตกต่างกันอย่างมีนัยสำคัญหรือไม่ โดยใช้สูตร

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

6

ในการจัดเป็นกลุ่มตัวอย่างคนละกลุ่ม

และใช้สูตร

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_m^2 + s_m^2 - 2rs_m s_m}}$$

7

ในการจัดเป็นกลุ่มตัวอย่างเดียวกัน

6. สหสัมพันธ์ของอำนาจจำแนก (Serial Correlation) เป็นการหาความสัมพันธ์ของอำนาจจำแนกของข้อมูลที่มีมากกว่า 2 ตัวแปรขึ้นไป ถ้าสัมประสิทธิ์สหสัมพันธ์สูง แสดงว่ามีอำนาจจำแนกของข้อสอบในการแยกเด็กเก่ง เด็กอ่อนได้ดี โดยใช้สูตร

$$r_p = \frac{\sum [(y_1 - y_h)]}{\sqrt{\sum [(y_1 - y_h)^2]}}$$

8

เมื่อ r_p แทน ค่าสหสัมพันธ์ของอำนาจจำแนก

y_1 แทน ค่าจุดตัดของแกน (ordinate) ที่มีค่าต่ำ

y_h แทน ค่าจุดตัดของแกนที่มีค่าสูง

⁶ Ibid. p. 191.

⁷ Ibid. p. 193.

⁸ Wert, James E., Neidt, Charles O., and Ahmann, J. Stanley, Statistical Method in Education and Psychological Research, p. 272.

$\bar{X}$ แทน คะแนนเฉลี่ยของข้อมูลแต่ละประเภท

6 แทน คะแนนเบี่ยงเบนมาตรฐาน

p แทน พหุคูณแต่ละประเภทของข้อมูล

7. สหสัมพันธ์พหุคูณของอำนาจจำแนก (Multiple Serial Correlation)

คือค่าสหสัมพันธ์เพื่อที่จะทราบว่าตัวพยากรณ์ ส่งผลต่ออำนาจจำแนกเพียงใด โดยใช้สูตร

$$R = \sqrt{\frac{\Delta}{N \sum \left[\frac{(y_l - y_h)^2}{p} \right]}}$$

เมื่อ R แทน สหสัมพันธ์พหุคูณของอำนาจจำแนก

Δ แทน ผลรวมของผลคูณระหว่าง สัมประสิทธิ์ของ serial correlation กับคะแนนเบี่ยงเบนจากคะแนนเฉลี่ย

N แทน จำนวนกลุ่มตัวอย่างทั้งหมด

⁹ Wert, Jame E., Neidt, Charles C; and Ahmann, J Stanley, op. cit. p.275.

ผลของการวิเคราะห์

ส่วนใหญ่ของบทนี้ จะเป็นการ เสนอผลวิเคราะห์เชิงสถิติของข้อมูลต่าง ๆ จากที่ได้ รวบรวมมา พร้อมกับการแปลความหมายของตัวเลขเหล่านั้นประกอบ. แต่อย่างไรก็ดี ในการ วิเคราะห์แต่ละตอนนั้น ก็จะกล่าวถึงที่มาและความเหมาะสมของเทคนิคสถิติที่ใช้ในการวิเคราะห์ นั้นไว้ด้วย. ส่วนรายละเอียดของเทคนิคและสูตรต่าง ๆ ที่นำมาใช้นั้น จะนำไปแสดงไว้ใน ภาคผนวก ก.

โครงร่างการวิเคราะห์ของรายงานนี้ แบ่งออกเป็น 3 หัวข้อใหญ่ และจะนำไปเสนอ เรียงไปตามลำดับ ดังนี้

ก. คุณภาพของแบบทดสอบในดานความยากง่าย (Difficulty) ในข้อนี้จะได้พยายาม เปรียบเทียบคะแนนเฉลี่ยจากผลการสอบของนักเรียนแต่ละประเภท ซึ่งจะแบ่งเป็นกลุ่มย่อย ๆ ตาม ระดับชั้นเรียน, ตามระดับอายุ, และตามเพศ. โดยจะทำการ เปรียบเทียบกันเป็นราย ๆ แบบ ทดสอบ. ผลที่ได้จากการวิเคราะห์และการทดสอบนี้สำคัญของความแตกต่างเหล่านี้ จะช่วยให้ เราสามารถประเมินความยากง่ายของแบบทดสอบแต่ละฉบับและแต่ละชนิดได้.

ข. คุณภาพของแบบทดสอบในดานอำนาจจำแนก (Discrimination) โดยนี้เกี่ยวข้องกับ ข้อแรก จะนำข้อมูลเหล่านั้นมาวิเคราะห์หาอำนาจจำแนกของแบบทดสอบแต่ละฉบับ โดยใช้ เทคนิคสถิติแบบ Discriminant Function ซึ่งเป็นขบวนการที่ค่อนข้างจะซับซ้อน แต่ก็ใช่วิธี ที่สามารถให้คำตอบที่สมบูรณ์และเที่ยงตรงมากกว่าเทคนิคสามัญแบบอื่น ๆ. ผลที่ได้จากการวิเคราะห์นี้ จะสามารถบอกให้เราทราบว่า แบบทดสอบแต่ละฉบับมีสมรรถภาพในการจำแนกแยะเด็กที่เขา สอบ ว่าควรจะเป็นนักเรียนอยู่ในกลุ่มชั้นใด หรือมีความสามารถอยู่ในระดับอายุและเพศใด เป็นต้น. อำนาจการจำแนกของแบบทดสอบนี้ เป็นกรรมวิธีที่สำคัญที่จะช่วยให้เราประเมินคุณภาพของแบบทดสอบ ได้ถูกต้องยิ่งขึ้น.

ค. สมการพยากรณ์สำหรับทำนายระดับชั้นเรียนและระดับอายุ ผลที่ได้จากการวิเคราะห์ของข้อนี้ จะออกมาเป็นรูปของสมการพยากรณ์ ว่าผู้สอบคนใดมีความสามารถเทียบเท่ากับนักเรียนชั้นใด หรือเทียบเสมือนกับความสามารถของนักเรียนอายุเท่าใด เป็นต้น.

ขบวนการวิเคราะห์ของเรื่องนี้ จะใช้วิธีสถิติแบบ Multiple serial correlation ซึ่งมีความซับซ้อนไม่น้อยกว่าข้อก่อน และเป็นเทคนิคที่ยังไม่เคยปรากฏว่าใครเคยใช้ในวงการวิจัยของเราเลย. จุดเด่นของวิธีวิเคราะห์นี้ อยู่ตรงที่มีความเหมาะสมกับลักษณะข้อมูลของรายงานนี้ และสามารถให้คำตอบที่เที่ยงตรงมากกว่าวิธีสถิติอื่น ๆ.

สัญลักษณ์ที่ใช้ในการวิเคราะห์

เนื่องจากการวิเคราะห์และการแปลผลที่จะกล่าวต่อไปนี้ เกี่ยวข้องกับตัวเลขและความหมายทางสถิติอยู่มาก จึงเห็นสมควรที่จะได้กำหนดสัญลักษณ์และอักษรย่อ เพื่อให้การเสนอนั้นเข้าใจง่ายและมีความหมายตรงกัน ดังนี้

- X แทน คะแนนดิบ (raw score or gross score) ที่นักเรียนสอบได้
- X_1 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ (Verbal Analogy)
- X_2 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ (Nonverbal Analogy)
- X_3 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ (Verbal Classification)
- X_4 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ (Nonverbal Classification)
- $\sum X$ แทน ผลรวมทั้งหมดของคะแนนดิบ
- $\sum X^2$ แทน ผลรวมของคะแนนดิบแต่ละตัวยกกำลังสอง
- $\bar{X}$ แทน คะแนนเฉลี่ย
- s แทน ความเบี่ยงเบนมาตรฐาน
- s^2 แทน ความแปรปรวน
- N แทน จำนวนนักเรียน
- r แทน สัมประสิทธิ์สหสัมพันธ์แบบ Product-moment correlation
- $r_{p.ser}$ แทน สหสัมพันธ์ของอำนาจจำแนก หรือ Point Serial correlation

$R_{p.ser}$ แทน สหสัมพันธ์พหุคูณของอำนาจจำแนก หรือ Multiple Serial correlation
 ρ แทน นำหนักของคะแนนมาตรฐาน
 B แทน นำหนักของคะแนนดิบ
 r_{tt} แทน ความเชื่อมั่นของแบบทดสอบ
 SE_{meas} แทน ความคลาดเคลื่อนมาตรฐานของคะแนน

ทั้งนี้จะได้นำเสนอผลวิเคราะห์ไปตามลำดับ ดังนี้

ก. คุณภาพของแบบทดสอบในด้านการความยากง่าย

การวิเคราะห์ในตอนนี้ จะเริ่มด้วยการเสนอ ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ จากคะแนนผลการสอบของนักเรียนทั้งหมดที่อยู่ในข่ายการทดลองจำนวน 1316 คน เพื่อจะได้นำมาพิจารณาคุณภาพโดยรวมของแบบทดสอบเหล่านั้นเสียก่อน. จากนั้นจึงจะได้นำมาวิเคราะห์เป็นรายฉบับ เป็นรายชั้นเรียน รายอายุ และเป็นรายเพศตามลำดับ. จุดมุ่งหมายสำคัญของการนี้ ก็เพื่อจะค้นหาให้แน่ชัดว่าแบบทดสอบฉบับใดมีความยากง่ายเหมาะสมกับนักเรียนระดับใดหรือกับกลุ่มใดมากที่สุด และการที่เรา นำแบบทดสอบชุดเดียวกันไปทดสอบกับนักเรียนต่างกลุ่มกันนั้น จะเกิดการแปรผันในเชิงสถิติ และในด้านการความยากง่ายเช่นไรด้วย. ทั้งนี้จะได้นำเสนอผลการวิจัยไปตามลำดับ ดังนี้

ก. 1 ค่าสถิติพื้นฐานของแบบทดสอบ

ในการวิจัยนี้ ผู้วิจัยได้พยายามวางแผนควบคุมตัวแปรที่คาดว่าจะส่งผลกระทบต่อคะแนนการสอบ และให้เป็นแบบเดียวกันหลายประการ. เช่นนอกจากจะกำหนดให้แบบทดสอบทั้งประเภทตัวหนังสือ และประเภทรูปภาพมีจำนวนข้อคำถามฉบับละ 40 ข้อเท่า ๆ กัน และให้เป็นคำถามที่ประกอบด้วยเนื้อหาเดียวกัน ถามในเรื่องเดียวกันแล้ว ยังได้กำหนดให้นักเรียนทุกคนและทุกชั้น ใช้เวลาในการทดสอบเพียงฉบับละ 20 นาทีเท่า ๆ กันหมด และนักเรียนทุกคนจะต้องสอบทั้ง 4 ฉบับเรียงไปตามลำดับก่อนหลังเป็นแบบเดียวกันอีกด้วย. การควบคุมทั้งนี้ โดยประสงค์ ก็เพื่อจะลดความคลาดเคลื่อนที่เป็นระบบ (systematic errors) ให้มีโอกาสดังกล่าวให้น้อยที่สุด อันจะส่งผลให้การทดลองมีความหมายยิ่งขึ้น ดังที่แสดงไว้ในตารางข้างล่างนี้.

ตาราง 5. ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ

แบบทดสอบ	จำนวน ข้อสอบ	N	$\bar{X}$	s^2	s	r_{tt}	$SE_{meas.}$
อุปมาอุปไมย-ตัวหนังสือ	40	1316	17.8093	54.4967	7.3822	.855	± 2.80811
อุปมาอุปไมย-ภาพ	40	1316	20.2659	47.1049	6.8633	.808	± 3.01243
ไม่เขาพวก-ตัวหนังสือ	40	1316	18.0327	51.5282	7.1783	.829	± 2.97268
ไม่เขาพวก-ภาพ	40	1316	22.8951	35.1372	5.9277	.740	± 3.02366

ตัวเลขในตารางนี้ขอนำสังเกตดูว่า ทั้ง ๆ ที่การวิจัยนี้พยายามควบคุมตัวแปรภายนอกให้เป็นเอกพันธ์แล้วก็ตาม แต่ผลการทดลองนี้ยังมีแนวโน้มชี้ให้เห็นว่า รายเฉลี่ยของแบบทดสอบประเภทที่ใช้ตัวหนังสือ จะต่ำกว่าประเภทรูปภาพอยู่เสมอ ดังที่แสดงไว้ในช่องคะแนนเฉลี่ยนั้นแล้ว และอีกประการหนึ่ง ค่าความแปรปรวนของแบบทดสอบประเภทตัวหนังสือ ก็จะมีปริมาณสูงกว่าประเภทรูปภาพอีกด้วย.

ค่าสถิติทั้ง 2 ชนิดนี้อาจแปล ความหมายรวมกันได้ว่า แบบทดสอบที่ใช้ตัวหนังสือเป็นคำถาม จะมีระดับความยากสูงกว่าประเภทรูปภาพสำหรับนักเรียนกลุ่มนี้เสมอ และในขณะเดียวกัน แบบทดสอบประเภทตัวหนังสือก็จะมีอำนาจการกระจายของคะแนนสูงกว่าด้วย. ปรากฏการณ์ชี้ให้เห็นชัดว่า เด็กนักเรียนในระดับชั้นประถมศึกษาจะมีสมรรถภาพในภาษา โดยเฉพาะอย่างยิ่งในด้านการอ่านตัวหนังสือ แตกต่างจากกันมาก, และจะต่างกับการอ่านรูปภาพอีกด้วย.

ในด้านความเชื่อมั่นของแบบทดสอบนั้น ก็ปรากฏผลดังในตารางข้างต้นนี้เช่นกันว่า ทั้งแบบทดสอบประเภทตัวหนังสือและประเภทรูปภาพ ต่างจะให้ค่าของความเชื่อมั่นเกือบไม่ผิดแผกไปจากกันเลย และแต่ละชนิดก็จะมีค่าความคลาดเคลื่อนมาตรฐานของคะแนนสอบต่างกันเล็กน้อยมาก เพียง 0.2 คะแนนเท่านั้นสำหรับแบบทดสอบอุปมาอุปไมย ($3.0124 - 2.8081$) และเพียง 0.05 คะแนนในแบบทดสอบไม่เขาพวก ($3.0237 - 2.9727$). นั่นคือ แบบทดสอบที่ใช้ตัวหนังสือหรือรูปภาพเป็นข้อคำถามนั้น ต่างจะมีคุณภาพในด้านความเชื่อมั่นของคะแนนเท่ากัน.

ก. 2 การทดสอบความแตกต่างของคะแนนเฉลี่ย

จากแนวโน้มความแตกต่างระหว่างแบบทดสอบทั้ง 2 ชนิดและ 2 ประเภทนั้น ทำให้ใคร่ทราบต่อไปอีกว่า ความแตกต่างของคะแนนเฉลี่ยเหล่านั้น จะมีความเชื่อถือในเชิงสถิติไ้มากน้อยเพียงใด. ในการนี้ จึงได้นำข้อมูลเหล่านั้นมาวิเคราะห์หาค่าความแปรปรวน (Analysis of Variance) แบบ 2×2 Factorial design ซึ่งก็ได้ผลการวิเคราะห์ดังตัวเลขในตาราง 6.

ตาราง 6. ผลการวิเคราะห์ความแปรปรวนของแบบทดสอบ 2 ชนิดทั้ง 2 ประเภท

แหล่งของความแปรปรวน	Sum of Squares	df	Mean Square	F	p
ระหว่างชนิดของแบบทดสอบ	2683.375	1	2683.375	14.2097	< .01
ระหว่างประเภทของแบบทดสอบ	17630.736	1	17630.736	93.3628	< .01
ระหว่างชนิด $\times$ ประเภทของแบบทดสอบ	1897.945	2	948.973	5.0252	< .05
ระหว่างภายใน	247571.166	1311	188.841		
รวมทั้งหมด	269783.222	1315			

สาเหตุที่การวิจัยนี้ใช้เทคนิค Analysis of Variance แบบ factorial design มาทำการวิเคราะห์ก็เพราะเห็นว่า ในประการแรกนั้น การทดลองนี้ได้นิยามควบคุมตัวแปรให้เป็นเอกเทศไว้แล้วหลายประการ จนอาจเชื่อได้ว่าแต่ละกลุ่มจะมีความคลาดเคลื่อนตามระบบเท่ากัน และความสำคัญประการที่สองนั้นอยู่ตรงที่การวิเคราะห์แบบนี้สามารถให้ค่าปฏิกริยารวมของความแปรปรวน (Interaction variance) ซึ่งในขั้นนี้ก็คิดหาค่ารวมระหว่างประเภทและชนิดของแบบทดสอบ ดังตัวเลขที่แสดงไว้ในบรรทัดที่สามของตารางนี้แล้ว.

ผลการวิเคราะห์ปรากฏว่า มีความแปรปรวนแตกต่างกันจริงในทุกลักษณะตัวแปร คือการที่แบบทดสอบใบข้อคำถามที่เป็นตัวหนังสือหรือรูปภาพนั้น จะมีผลทำให้ความแปรปรวนของคะแนนเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญเชิงสถิติอย่างแน่นอน. ยิ่งเป็นแบบทดสอบที่วัดตามสมรรถภาพคนละชนิด เช่น จินตภาพมาอุปไมยและชนิดไม่เขาพวกควายแล้ว ก็จะมีผลต่อความแปรปรวนของคะแนนเฉลี่ยมากขึ้นด้วย. ส่วนความสำคัญประการสุดท้ายของการวิเคราะห์ปฏิกริยารวม (Interaction) ที่มีความสูงจนมีนัยสำคัญเชิงสถิตินั้น ก็มีความหมายว่า การที่เด็กนักเรียนจะสอบได้คะแนนมากหรือน้อยขึ้นอยู่กับชนิดและประเภทของแบบทดสอบเป็นประการสำคัญ มิใช่ว่าใครเก่งในแบบทดสอบชนิดใดแล้วจะสามารถสอบได้สูงในแบบทดสอบชนิดอื่น ๆ หรือประเภทอื่น ๆ ด้วย ก็หาไม่ได้. นั่นคือ ทั้งประเภทและชนิดของแบบทดสอบต่างจะมีอิทธิพลต่อคะแนนผู้สอบเป็นราย ๆ ลักษณะไป.

ผลการวิเคราะห์แบบ Factorial design ดังกล่าวมาแล้วนี้ ยังเป็นการเปรียบเทียบความแตกต่างเป็นส่วนรวมระหว่างนักเรียนทั้ง 3 ระดับ กับแบบทดสอบทั้ง 4 ฉบับอยู่ ไม่สามารถชี้ให้เห็นชัดลงไปว่า แบบทดสอบชนิดใดหรือประเภทใดที่มีคะแนนเฉลี่ยแตกต่างจากกันและกันจริง. จึงเห็นสมควรที่จะได้นำคะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับนั้นมาเปรียบเทียบสลับกันไปทีละคู่ ๆ จนครบ 6 คู่ ว่าคู่ใดบ้างที่มีคะแนนเฉลี่ยสูงกว่ากันจริงในเชิงสถิติ ดังที่แสดงไว้ในตารางที่ 7.

ในการนี้ ได้นำผลต่างของคะแนนเฉลี่ยของแบบทดสอบแต่ละคู่ มาทดสอบนัยสำคัญด้วยวิธี t -test ดังสูตรที่แสดงไว้แล้วในภาคผนวก ก. ผลการทดสอบปรากฏว่า คะแนนเฉลี่ยของแบบทดสอบคู่ใดมีความแตกต่างกันอย่างมีนัยสำคัญ เราก็อาจสรุปได้ว่า แบบทดสอบคู่หนึ่งมีความยากง่ายแตกต่างจากกันจริงจึงมั่นใจได้. และเพื่อให้การเสนอผลเห็นง่าย จึงได้แสดงผลต่างของคะแนนเฉลี่ยไว้ในบรรทัดบน และแสดงค่าของ t ไว้ภายในวงเล็บแนบด้วย. ถ้าความแตกต่างนั้นมีมากจนอาจมั่นใจได้ถึงระดับ 5 % แล้ว ก็จะเติมดอกจันไว้เหนือตัวเลขนั้นหนึ่งดอก และถ้าคู่ใดแตกต่างกันมาก จนอาจเชื่อมั่นได้ถึงระดับ 1 % หรือเกินกว่า ก็จะใส่ดอกจันไว้สองดอก ส่วนคู่ที่มีความแตกต่างกันน้อยจนไม่เป็นที่มั่นใจ ก็จะไม่ใส่ดอกจันไว้เลย.

ตาราง 7. ผลต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ และระดับนัยสำคัญของผลต่างเหล่านั้น

แบบทดสอบ	อุปมาอุปไมย - ตัวหนังสือ	อุปมาอุปไมย - ภาพ	ไม่เขาพวก - ตัวหนังสือ	ไม่เขาพวก - ภาพ
อุปมาอุปไมย-ตัวหนังสือ	—	2.4566 ^{**} (t = 26.90)	.2234 ^{**} (t = 5.77)	5.0858 ^{**} (t = 22.52)
อุปมาอุปไมย-ภาพ		—	2.2332 ^{**} (t = 13.88)	2.6292 ^{**} (t = 22.63)
ไม่เขาพวก-ตัวหนังสือ			—	4.8624 ^{**} (t = 36.23)
ไม่เขาพวก-ภาพ				—

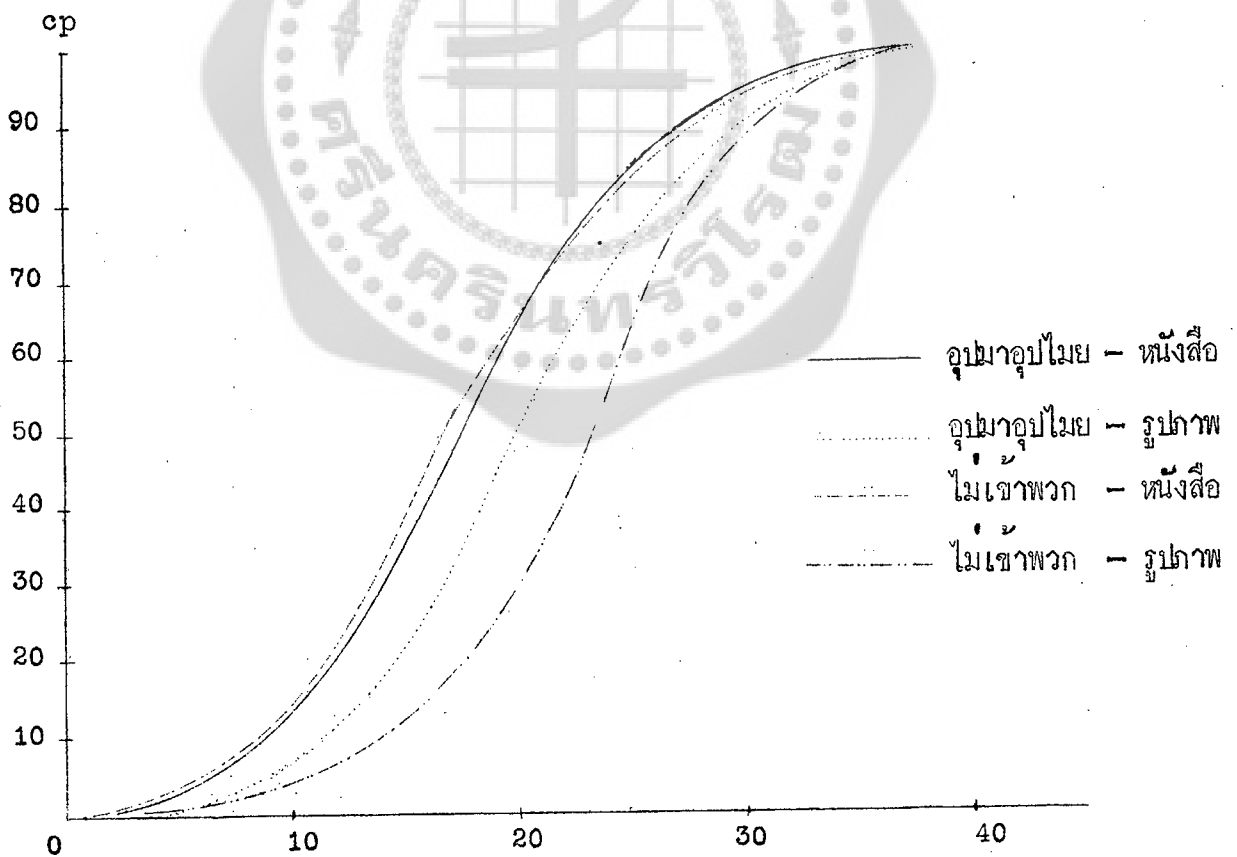
ผลการทดสอบด้วย t - test ปรากฏว่า แบบทดสอบทั้ง 2 ชนิดและ 2 ประเภท ต่างจะมีความแตกต่างในระดับความยากง่ายกว่ากันและกันจริงหมดทุกคู่ ณ ระดับความเชื่อมั่นเกินกว่า 1 % ซึ่งก็หมายความว่า ถึงจะทำการทดลองแบบนัย ๆ กันถึงร้อยครั้งก็จะให้ผลเกินกว่า 99 ครั้งว่าแต่ละฉบับจะให้คะแนนมากและน้อยผิดแผกจากกันเป็นทำนองเดียวกันกับที่ทดลองในครั้งนี้.

อนึ่ง จากการพิจารณาของ t ในตารางที่ 7. ก็จะได้ความจริงอีกประการหนึ่งว่า ในแบบทดสอบชนิดเดียวกันที่ใช้คำถามคนละประเภท จะยังมีคะแนนเฉลี่ยแตกต่างกัน หรืออีกคือความยากง่ายผิดแผกจากกัน มากกว่าแบบทดสอบคู่อื่น ๆ เสมอ. เช่น ในแบบทดสอบอุปมาอุปไมยที่ใช้คำถามเป็นตัวหนังสือกับรูปภาพนั้น จะมีค่า t สูงมากถึง 26.90 และโดยทำนองเดียวกัน ในแบบทดสอบไม่เขาพวกที่ใช้คำถามเป็นตัวหนังสือและรูปภาพนั้น ก็จะมีค่า t ยิ่งสูงถึง 36.23 ด้วย. ส่วนแบบทดสอบอุปมาอุปไมยและไม่เขาพวกที่เป็นตัวหนังสือด้วยกันทั้งคู่ จะมีค่า t เพียง 5.77 เท่านั้น เป็นต้น.

ฉะนั้น ในขั้นนี้เราจึงอาจลงสรุปได้อย่างมั่นใจว่า แบบทดสอบที่ใช้คำถามเป็นตัวหนังสือ และรูปภาพนั้น ต่างจะมีอิทธิพลต่อคะแนนสอบของนักเรียนระดับชั้นประถมศึกษาโดยไม่คงสงสัย. และในระหว่างแบบทดสอบทั้ง 2 ประเภทนี้ ประเภทที่ใช้รูปภาพจะมีระดับความยากน้อยกว่า หรือก็อาจง่ายกว่าประเภทที่ใช้ตัวหนังสือเสมอ.

ผลวิเคราะห์ประการแรกที่ได้นี้ อาจใช้เป็นข้อเตือนใจนักการศึกษาทั่ว ๆ ไป โดยเฉพาะอย่างยิ่งกับผู้ที่เกี่ยวข้องในวงการวัดระดับประถมศึกษาว่า สมควรจะใส่ใจแบบทดสอบประเภทใด ในการวัดผลสัมฤทธิ์ทางการศึกษา อันเป็นข้อเท็จจริงที่วงการศึกษามีอาจมองข้ามเสียได้เลย.

เพื่อเสนอผลของการเปรียบเทียบดังกล่าวข้างต้นให้เห็นชัดยิ่งขึ้น จึงได้นำการแจกแจงคะแนนของแบบทดสอบทั้ง 4 ฉบับนั้นมาสร้างเป็นเส้นโค้ง ogive ดังแสดงไว้ในภาพที่ 1 ต่อไปนี้.



ภาพที่ 1. ภาพการ เปรียบเทียบความสามารถของนักเรียนชั้นประถมศึกษาในแต่ละวิชา

เส้น ogive ทั้ง 4 ให้ความหมายเชิงเปรียบเทียบแก่เราเป็นอย่างดีว่า ในด้านความยากง่ายนั้น แบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือจะมีความยากมากที่สุด ฅบที่ยากรองลงมา ก็ไคแกแบบทดสอบไมเซาพวกประเภทตัวหนังสือ, อุปมาอุปไมยประเภทรูปภาพ, และไมเซาพวกประเภทรูปภาพตามลำดับ. เรืองนันทนาสังเกทว่า แบบทดสอบประเภทที่เป็นตัวหนังสือทั้ง 2 ฅบ จะยากกว่าประเภทรูปภาพอีก 2 ฅบเสมอ.

อนึ่ง ถ้าไคสังเกทความลาดชัน (slope) ของเส้นไคทั้ง 4 เส้น ก็จะได้เห็นอีกเช่นกันว่า ogive 2 เส้นหลังของแบบทดสอบประเภทรูปภาพ มีไคไครงไคไลการแจกแจงแบบไคปกติ (Normal curve) มากกว่า ogive ของประเภทหนังสือมาก. ฉะนั้น ถ้าเราก่อหลักการว่า ฅบแบบทดสอบที่ไคจะต้องให้คะแนนแจกแจงเป็นปกติแล้ว เราก็อาจประเมินค่าแบบทดสอบเหล่านี้ โดยสรุปไคได้ว่า ในนักเรียนระดับชั้นประถมศึกษา แบบทดสอบประเภทรูปภาพจะมีคุณค่าในเชิง วิชาผลมากกว่าแบบทดสอบที่ใชตัวหนังสือ.

ก. 3 คุณภาพของแบบทดสอบจำแนกเป็นรายชั้นเรียน

การศึกษาในตอนนี้จะไคนำข้อมูลทั้งหมดมาแยกทำการวิเคราะห์เป็นราย ๆ ชั้น. ทั้งนี้ โดยประสงค์เพื่อจะสำรวจว่า คุณภาพเชิงสถิติของแบบทดสอบเหล่านี้จะเปลี่ยนแปลงเช่นไร เมื่อนำไปใช้กับนักเรียนในระดับชั้นประถมปีที่ 2, ปีที่ 4, และปีที่ 7.

วิธีดำเนินการของเรื่องนี้ ก็โดยจำแนกข้อมูลทั้ง 1316 คน ออกเป็นรายชั้นเรียน แล้วคำนวณหาสถิติต่าง ๆ ของแต่ละชั้นเรียนเหล่านี้แบบเดียวกับที่กระทำมาแล้วในข้อ ก. 1 ซึ่งไคไคผล ดังในตารางที่ 8. ต่อไปนี้

อนึ่ง ในการวิจัยนี้ ไคนำตัวเลขคะแนนดิบของแต่ละแบบทดสอบ จากนักเรียนทั้งหมด จำนวน 1316 คน มาแจกแจงความถี่ แล้วไคทำการแปลงรูปคะแนนดิบเหล่านั้นให้เป็นคะแนน T - ปกติ (Normalized T Score by Area Transformation) ซึ่งเป็นคะแนนรูปใหม่ที่มีหน่วย เท่ากันทุกตัว และสามารถนำมาเปรียบเทียบกันได้โดยตรงอย่างมีความหมาย. บัญชีการ แปลงคะแนนดิบออกเป็นคะแนน T - ปกตินี้จะไคแสดงไว้ในภาคผนวก ข. และจะไคใช้เป็นเกณฑ์ ปกติ (Norm) สำหรับการวิจัยต่อไป.

ตาราง 8. การศึกษาพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ จำแนกเป็นรายชั้นเรียน

ชั้น แบบทดสอบ	ประดมปีที่ 2	ประดมปีที่ 4	ประดมปีที่ 7
<u>อุปมาอุปไมย-ตัวหนังสือ</u>			
N	424	467	425
$\bar{X}$	11.8997 = T43	16.7516 = T50	24.8871 = T58
s^2	16.7159	32.4145	30.1098
s	4.0885	5.6934	5.4872
r_{tt}	.5132	.7175	.7053
$SE_{meas.}$	± 2.8525	± 3.0260	± 2.9784
<u>อุปมาอุปไมย-ภาพ</u>			
N	424	467	425
$\bar{X}$	14.4198 = T42	19.8073 = T50	26.6024 = T58
s^2	15.2512	26.3447	27.2683
s	3.9053	5.1327	5.2219
r_{tt}	.4054	.6363	.6905
$SE_{meas.}$	± 3.0113	± 3.0950	± 2.9049
<u>ไม่เข้าพวก-ตัวหนังสือ</u>			
N	424	467	425
$\bar{X}$	11.9316 = T42	17.3640 = T50	24.8541 = T58
s^2	16.6832	32.2963	23.3135
s	4.0845	5.6830	4.8284
r_{tt}	.5109	.7135	.6116
$SE_{meas.}$	± 2.8562	± 3.0415	± 3.0090
<u>ไม่เข้าพวก-ภาพ</u>			
N	424	467	425
$\bar{X}$	18.2028 = T42	23.1799 = T49	27.2635 = T57
s^2	26.7079	23.4246	16.5907
s	5.1680	4.8399	4.0732
r_{tt}	.6218	.5943	.4889
$SE_{meas.}$	± 3.1778	± 3.0825	± 2.9119

ตารางนี้ให้ความหมายเชิงสถิติของแบบทดสอบใดหลายประการ ดังจะได้เสนอเป็นข้อ ๆ ดังนี้

1. โดยทั่วไปคะแนนรายเฉลี่ยของแบบทดสอบแต่ละฉบับจากชั้นประถมศึกษาปีที่ 2 - 7 จะมีค่าเพิ่มสูงขึ้นเรื่อย ๆ ตามลำดับชั้นเรียน ซึ่งก็มีเหตุผลสมควรที่จะเป็นไปตามนั้น เพราะนักเรียนมีความมั่งคั่งทางพัฒนาสูงขึ้นกว่าเดิม. แต่สิ่งที่สำคัญกว่าของเรื่องนี้อยู่ตรงที่ปริมาณการเพิ่มของคะแนนจากระดับหนึ่งไปสู่อีกระดับหนึ่ง ซึ่งมีความหมายแทนอัตราการงอกงามของนักเรียนเหล่านั้น. เพื่อจะทราบถึงการเปลี่ยนแปลงนี้ จึงได้แปลงคะแนนเฉลี่ยของทุกวิชาและทุกชั้นไปเป็นคะแนน T ประกติ ดังที่แสดงไว้ในตารางที่ 8. นั้นด้วย ซึ่งปรากฏผลดังนี้

แบบทดสอบ	ปริมาณที่เพิ่มจาก ป.2 ถึง ป.4		ปริมาณที่เพิ่มจาก ป.4 ถึง ป.7	
อุปมาอุปไมย-ตัวหนังสือ	จาก T43 เป็น	T50 = +7	จาก T50 เป็น	T58 = +8
อุปมาอุปไมย-ภาพ	จาก T42 เป็น	T50 = +8	จาก T50 เป็น	T58 = +8
ไม่เขาพวก-ตัวหนังสือ	จาก T42 เป็น	T50 = +8	จาก T50 เป็น	T58 = +8
ไม่เขาพวก-ภาพ	จาก T42 เป็น	T49 = +7	จาก T49 เป็น	T57 = +8

ตัวเลขที่หน้าตั้งเกิดอยู่ 2 ประการ. ในประการแรกนั้นจะเห็นว่า แบบทดสอบทั้ง 4 ฉบับนี้มีความยากง่ายเหมาะสมกับนักเรียนชั้นประถมศึกษาปีที่ 4 มากที่สุด เพราะคะแนนเฉลี่ยของทุกฉบับมีค่าเท่ากับหรือใกล้เคียงกับคะแนน T 50 ทั้งสิ้น ส่วนคะแนนเฉลี่ยทุกฉบับของชั้นประถมศึกษาปีที่ 2 จะได้เพียงคะแนน T 42 เท่านั้น แปลว่ายากเกินไปสำหรับเด็กชั้นนี้. ส่วนสำหรับชั้นประถมศึกษาปีที่ 7 นั้นแบบทดสอบเหล่านี้จะง่ายเกินไป เพราะตรงกับคะแนน T ที่ 58.

สิ่งที่น่าสังเกตประการที่สองนั้น ก็ได้แก่ปริมาณการเพิ่มขึ้นของคะแนนเฉลี่ยในแต่ละแบบทดสอบจากระดับชั้นประถมศึกษาปีที่ 2 ถึง ชั้นประถมศึกษาปีที่ 4 ซึ่งจะมีอัตราการเพิ่มประมาณ 7 หรือ 8 หน่วย เท่า ๆ กับปริมาณที่เพิ่มจากประถมศึกษาปีที่ 4 ถึง ประถมปีที่ 7 ที่เพิ่มขึ้น 8 หน่วยเท่ากันหมด. เมื่อมาคำนึงว่า จากชั้นประถมศึกษาปีที่ 2 ถึง ประถมปีที่ 7 เป็นระยะพัฒนาเพียง 2 ปี ส่วนจากประถมศึกษาปีที่ 4 ถึง ประถมปีที่ 7 นั้นมีเวลา 3 ปี เราจึงอาจอนุมานได้ว่า เด็กในระหว่างอายุ 9 - 11 ขวบ จะมีอัตราการพัฒนาได้มากปาน ๆ กันระหว่างอายุ 11 ถึง 14 ขวบทีเดียว.

ปรากฏการณ์อาจมีความสัมพันธ์กับกราฟความงอกงาม (growth curve) ของสิ่งมีชีวิต ดังที่ทราบกันโดยทั่ว ๆ ไปว่า จะมีอัตราเพิ่มขึ้นอย่างรวดเร็วในระยะแรก แล้วจะคงที่ในตอนกลาง และค่อย ๆ ลดลงในขั้นปลาย ก็อาจเป็นได้.

2. ค่าความแปรปรวน (variance) ของคะแนนในกลุ่มนักเรียนเหล่านี้ ก็มีแนวโน้มทางเพิ่มมากขึ้นตามระดับชั้นเรียนที่สูงขึ้น. เรื่องนี้อาจหมายความว่า เมื่อนักเรียนยังอยู่ในชั้นต่ำ และมีอายุอ่อนนั้นต่างก็จะมีความสามารถคล้ายคลึงกัน ๆ กันเป็นส่วนใหญ่ ต่อเมื่อใช้เวลาเรียนนานปีขึ้นและได้รับประสบการณ์เพิ่มมากขึ้น ต่างจะสามารถฝึกแยกจากกันมากกว่าเมื่อครั้งยังเป็นเด็กเล็กอยู่ และเรื่องนี้ก็น่าสังเกตว่า ความแปรปรวนของคะแนนจะมีมากที่สุดในระดับชั้นประถมปีที่ 4 โดยเฉพาะอย่างยิ่งในแบบทดสอบอุปมาอุปไมยและไม่เข้าพวกประเภทตัวหนังสือ ซึ่งเป็นข้อสนับสนุนผลวิเคราะห์ที่กล่าวมาแล้วในหน้า 37 อีกครั้งหนึ่ง.

3. ในด้านความเชื่อมั่นของคะแนน (r_{tt}) นั้น ก็ปรากฏว่า แบบทดสอบชนิดอุปมาอุปไมยทั้ง 2 ประเภท จะมีค่าความเชื่อมั่นที่สูงขึ้นตามระดับชั้น คือยิ่งชั้นสูงขึ้นความแน่นอนของคะแนนจากแบบทดสอบก็จะยิ่งสูงขึ้นด้วย เป็นสำคัญว่า แบบทดสอบทั้ง 2 ฉบับนี้ มีคุณค่าเหมาะสมกับเด็ก ประถมปีที่ 7 มากที่สุด ส่วนแบบทดสอบไม่เข้าพวกทั้ง 2 ประเภท มีความเชื่อมั่นสูงเป็นรองลงมา โดยเฉพาะแบบทดสอบไม่เข้าพวกประเภทภาพ จะมีความเชื่อมั่นลดน้อยกว่าเมื่อใช้กับชั้นสูง ๆ ทั้งนี้เพราะแบบทดสอบฉบับนี้กลายเป็นแบบทดสอบที่ง่ายเกินไปสำหรับเด็กชั้นประถมปีที่ 7 จนเด็กชั้นนี้สามารถทำคะแนนเฉลี่ยได้สูงเกือบ 70 % ของข้อคำถามทั้งหมด.

4. ถ้าจะประเมินคุณค่าของแบบทดสอบเหล่านี้โดยส่วนรวม และโดยยึดหลักการวัดผลที่ว่าแบบทดสอบที่ดีต้องให้คะแนนเฉลี่ยใกล้เคียงกับครึ่งหนึ่งของจำนวนข้อคำถามแล้วจะปรากฏว่า แบบทดสอบที่เป็นตัวหนังสือจะให้คะแนนเฉลี่ยน้อยกว่าแบบทดสอบที่เป็นรูปภาพในทุกระดับชั้นเรียน. และสำหรับเด็กชั้นประถมปีที่ 2 นั้น แบบทดสอบที่เหมาะสมแก่เขามากที่สุดคือ แบบทดสอบประเภทภาพ ซึ่งจะให้คะแนนเฉลี่ยใกล้เคียงกับ 20 คะแนนมากกว่าแบบทดสอบประเภทตัวหนังสือเป็นอันมาก.

สำหรับนักเรียนชั้นประถมปีที่ 4 ก็เช่นเดียวกัน ผลวิเคราะห์ในตารางข้างต้นชี้ชัดว่าแบบทดสอบประเภทรูปภาพจะเหมาะสมกับเด็กในระดับนี้มาก โดยเฉพาะอย่างยิ่งชนิดอุปมาอุปไมยประเภทภาพ ซึ่งให้คะแนนเฉลี่ยถึง 19.81 ใกล้เคียงกับครึ่งหนึ่งของคะแนนเต็มมากที่สุด.

ตรงข้ามกับอุปมาอุปไมยที่เป็นภาษา ซึ่งกลายเป็นข้อสอบที่ยากเกินไปสำหรับเด็กในระดับนี้.

ส่วนในชั้นประถมศึกษาปีที่ 7 นั้น คะแนนเฉลี่ยของแบบทดสอบทั้งที่เป็นตัวหนังสือ และที่เป็นรูปภาพ ต่างได้คะแนนเกินครึ่งหนึ่งของคะแนนเต็มของแบบทดสอบทุกฉบับ ซึ่งก็แสดงว่า นักเรียนชั้นประถมศึกษาปีที่ 7 เป็นเด็กที่เข้าใจภาษาได้พอสมควรแล้ว ถึงจะใช่แบบทดสอบที่เป็นตัวหนังสือหรือรูปภาพก็จะให้ผลไม่ผิดแผกจากกันมากนัก.

ก. 4 การทดสอบความแตกต่างของคะแนนเฉลี่ยจำแนกตามระดับชั้นเรียน

จากผลวิเคราะห์ทางสถิติแสดงว่า แบบทดสอบบางฉบับจะมีระดับความยากง่ายเหมาะสมกับนักเรียนบางชั้นนั้น ทำให้ต้องการทราบต่อไปว่า จะมีโอกาสบางหรือไม่ที่นักเรียนชั้นใดจะมีความสามารถทำแบบทดสอบฉบับใดได้ดีหรือเลวเป็นพิเศษกว่าฉบับอื่น ๆ. คำตอบของปัญหานี้คือการทดสอบัยสำคัญของปฏิกริยารวม (Interaction variance) ระหว่างแบบทดสอบ x ชั้นนั่นเอง.

ฉะนั้น จึงนำข้อมูลเหล่านั้นมาวิเคราะห์ความแปรปรวน ดังแสดงไว้ในตารางที่ 9.

ตาราง 9. ผลการวิเคราะห์ความแปรปรวนของคะแนนสอบทั้ง 4 ฉบับ จำแนกตามระดับชั้น ประถมปีที่ 2, ประถมปีที่ 4 และประถมปีที่ 7

Source of Variation	Sum of Squares	df	Mean Square	F	p
Between tests	22212.056	3	7404.0186	75.1729	< .01
Between grades	118739.572	2	5936.9786	602.7811	< .01
Interaction : test x grades	396.509	6	66.0848	.6709	> .05
Within	128435.085	1304	98.4931		
Total	269703.222	1315			

ผลการวิเคราะห์ความแปรปรวนระหว่างแบบทดสอบ (Between tests) และระหว่างชั้น (Between grades) ปรากฏว่า มีความแปรปรวนต่างกันอย่างมีนัยสำคัญเชิงสถิติ ระดับน้อยกว่า .01 ทั้ง 2 ประการ ซึ่งก็หมายความว่า คะแนนที่ได้จากการทดสอบแต่ละฉบับมีความแตกต่างกันจริง และคะแนนของนักเรียนแต่ละชั้นก็แตกต่างกันอย่างมีนัยสำคัญด้วย. นัยสำคัญนี้เป็นสัญญาณให้เราอนุมานได้ว่า เด็กแต่ละชั้นมีความสามารถแตกต่างกันจริง ก็นักเรียนชั้นสูงจะได้คะแนนมากกว่าชั้นต่ำเสมอ. ส่วนปฏิกริยารวมของความแปรปรวน (Interaction variance) นั้น ในกรณีนี้ยังมีความแปรปรวนน้อยไม่ถึงระดับความเชื่อมั่นได้ ซึ่งก็หมายความว่า นักเรียนบางชั้นอาจมีความสามารถสูงหรือต่ำในแบบทดสอบบางฉบับก็ได้ ไม่มีความแน่นอนว่า นักเรียนชั้นใดจะมีความสามารถเด่นตอเป็นพิเศษ ในแบบทดสอบชนิดใดๆ โดยเฉพาะเลย ความแตกต่างของคะแนนที่สังเกตเห็นนั้นเป็นไปตามโชค (chance) โดยปราศจากระบบใด ๆ ทั้งสิ้น.

จากนั้น ได้นำคะแนนเฉลี่ยของแต่ละแบบทดสอบมา เปรียบเทียบกันเป็นราย ๆ ชั้นเรียนด้วยวิธี t - test ซึ่งนักเพื่อจะทดสอบดูว่า คะแนนเหล่านั้นจะแตกต่างกันและกัน อย่างมีนัยสำคัญเชิงสถิติหรือไม่ ดังที่แสดงไว้ในตารางที่ 10.

ตาราง 10. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ ในแต่ละ
ชั้นเรียน และระดับความเชื่อมั่น ของผลทางเหล่านี้น

	ชั้น ป.2	ชั้น ป.4	ชั้น ป.7
<u>อุปมาอุปไมย-ตัวหนังสือ</u>			
ชั้น ป.2	—	4.8519 ^{**} (t = 14.77)	12.9874 ^{**} (t = 39.17)
ชั้น ป.4		—	8.1355 ^{**} (t = 21.73)
<u>อุปมาอุปไมย-ภาพ</u>			
ชั้น ป.2	—	5.3875 ^{**} (t = 17.72)	12.1826 ^{**} (t = 38.50)
ชั้น ป.4		—	6.7951 ^{**} (t = 19.57)
<u>ไม่เข้าพวก-ตัวหนังสือ</u>			
ชั้น ป.2	—	5.4324 ^{**} (t = 16.49)	12.9225 ^{**} (t = 42.11)
ชั้น ป.4		—	7.4901 ^{**} (t = 21.27)
<u>ไม่เข้าพวก-ภาพ</u>			
ชั้น ป.2	—	4.9771 ^{**} (t = 15.03)	4.0836 ^{**} (t = 28.87)
ชั้น ป.4		—	4.0836 ^{**} (t = 13.67)

ผลการทดสอบปรากฏว่า คะแนนเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับในแต่ละชั้นเรียนมีความแตกต่างกันอย่างมีนัยสำคัญอย่างเชื่อมั่นได้เกินระดับ .01 ทั้งสิ้น ซึ่งหมายความว่า แบบทดสอบเหล่านี้มีอำนาจจำแนกความสามารถของเด็กในแต่ละชั้นเรียนได้เป็นที่น่าพอใจ.

ก. 4 คุณภาพของแบบทดสอบฉบับนี้อุปมาอุปไมยที่เป็นตัวหนังสือจำแนกเป็นรายอายุ

ในตอนนี้จะได้ทำการศึกษาถึงคุณภาพของแบบทดสอบทั้ง 4 ชนิด ว่าจะมีการแปรผันไปตามระดับอายุของผู้สอบหรือไม่. ในการนี้ได้นำข้อมูลของนักเรียนทั้ง 1316 คน มาจำแนกเป็นกลุ่มย่อย ๆ ตามรายอายุ โดยใช้เกณฑ์ว่า จากอายุน้อยกว่า 7 ปี จนถึง 7 ปี 6 เดือน ให้นับเป็นกลุ่มอายุ 7 ปี. และในทำนองเดียวกัน จากอายุ 7 ปี 7 เดือน จนถึง 8 ปี 6 เดือน ก็จัดอยู่ในกลุ่มอายุ 8 ปี เป็นต้น.

ควมการจำแนกกลุ่มอายุตามวิธีนี้ จะสามารถแยกนักเรียนออกเป็น 9 กลุ่ม จากอายุ 7 ปี ถึง 15 ปี. จากนั้นได้นำข้อมูลของแต่ละกลุ่มอายุมาคำนวณหาค่ารายเฉลี่ย ($\bar{X}$) และได้แปลี่ยนคะแนนให้เป็นรูปของคะแนน T ประกติ ซึ่งมีความหมายแทนความสามารถปานกลางของแต่ละกลุ่มอายุนั้น ๆ และในฐานะที่คะแนน T ประกตินี้หน่วยเท่ากัน เราจึงอาจใช้คะแนนแปลงรูปนี้ สำหรับติดตามปริมาณความงอกงามของแต่ละอายุ ๆ ได้สะดวกยิ่งขึ้น.

นอกจากจะหาค่ารายเฉลี่ยแล้ว ยังได้นำข้อมูลของแต่ละอายุมาคำนวณหาความแปรปรวนและความเบี่ยงเบนมาตรฐาน (s^2, s) ซึ่งเป็นค่าสถิติแสดงปริมาณการกระจายของข้อมูลเหล่านั้น จากนั้นได้หาค่าความเชื่อมั่น (r_{tt}) ของแบบทดสอบจำแนกตามระดับอายุ พร้อมกับความคลาดเคลื่อนมาตรฐานของการวัด ($SE_{meas.}$) ประกอบกับ ค่าสถิติ 2 ประการหลังนี้ จะช่วยให้เราเห็นคุณภาพของแบบทดสอบแต่ละฉบับว่า เมื่อนำไปใช้กับนักเรียนที่มีอายุต่าง ๆ กันแล้ว จะเกิดการแปรผันเช่นไรหรือไม่. นั่นก็คือเพื่อจะค้นหาว่า อายุมีอิทธิพลต่อแบบทดสอบเหล่านี้หรือไม่.

ขบวนการวิเคราะห์ที่กล่าวมานี้ โดยเนื้อแท้ก็คือ ความพยายามทำนายเอาหลักการ เรื่องการสร้างเกณฑ์ปรกติอายุ (Age norm) แบบ Cross sectional study มาดัดแปลงให้เหมาะสมกับงานวิจัยนี้ ดังจะได้เสนอผลการวิเคราะห์พร้อมกับการแปลผลและทดสอบนัยสำคัญไปตามลำดับขั้นตอน

แบบทดสอบ จากอุปมาอุปไมยประเภทตัวหนังสือและรูปภาพ กับแบบทดสอบไม่เข้าพวกประเภท
ตัวหนังสือและรูปภาพ ดังนี้

ตาราง 11. ค่าสถิติพื้นฐานของแบบทดสอบอุปมาอุปไมย-ตัวหนังสือ จำแนกเป็นรายอายุ

ระดับอายุเป็นปี	N	$\bar{X}$	T-score	s^2	s	r_{tt}	$SE_{meas.}$
7	74	11.7297	41	19.4601	4.4114	.589	± 2.8290
8	154	12.2792	43	16.4378	4.0544	.495	± 2.8823
9	157	13.8408	45	30.9552	5.5637	.726	± 2.9133
10	208	15.8510	48	39.2965	6.2687	.776	± 2.9676
11	189	16.3651	49	29.1586	5.3999	.686	± 3.0283
12	157	19.9299	52	59.5656	7.7179	.853	± 2.9506
13	211	23.7867	56	41.5209	6.4281	.786	± 2.9711
14	108	24.7963	57	34.0186	5.8325	.741	± 2.9658
15	58	23.0862	56	36.9574	6.0793	.755	± 3.0105

ตารางนี้แสดงความจริงที่น่าสนใจอยู่ 5 ประการคือ

1. คะแนนเฉลี่ยของนักเรียนแต่ละระดับอายุมีแนวโน้มเพิ่มสูงขึ้นเรื่อย ๆ ตามลำดับอายุ
ของนักเรียน คือกลุ่มนักเรียนที่มีอายุมากจะมีความสามารถทำแบบทดสอบนี้ได้ดีกว่ากลุ่มที่มีอายุ
ยังน้อย. ปรากฏการณ์นี้เป็นที่ปรารถนาของผู้สร้างแบบทดสอบทั่ว ๆ ไปอย่างหนึ่ง เพราะอาจใช้เป็น
เป็นหลักฐานสนับสนุน ว่าแบบทดสอบนี้มีคุณภาพสูงและมีความไวในการจำแนกสมรรถภาพของนักเรียน
แต่ละระดับได้.

อนึ่ง ถ้าได้สังเกตค่าของคะแนน T ของแต่ละระดับอายุครบแล้ว ก็จะยิ่งเห็นชัดว่า
มีอัตราที่เพิ่มขึ้นค่อนข้างจะเป็นไปตามโค้งความงอกงาม (Growth curve) ครบ จากอัตรา

การเพิ่มประมาณ 2 หน่วย ในระหว่างอายุ 7 ถึง 8 และจาก 8 ถึง 9 ขวบ จากนั้น จะเพิ่มขึ้นเป็น 3 หน่วยในระหว่างอายุ 9 ถึง 10 และ 11 ถึง 12 ขวบ ส่วนจากอายุ 12 ถึง 13 ขวบจะเพิ่มถึง 4 หน่วย เป็นต้น. สำหรับของอายุจาก 14 ถึง 15 ขวบที่ไคเคแน T ลดลง 1 หน่วยนั้นเมื่อไควิเกราะที่มา ของข้อมูลแล้วปรากฏว่า ส่วนใหญ่เป็นกลุ่มนักเรียนอ่อนที่เรียนช้ากว่าปกติ และมักจะเป็นนักเรียนที่เคยสอบตกเรียนซ้ำชั้นเป็นส่วนมาก. แต่อย่างไรก็ดี ความคลาดเคลื่อนนี้ก็ป็นจำนวนน้อยมากไม่มีผลกระทบกระเทือนต่อการวิจัยนี้.

2. เป็นที่น่าสนใจอีกประการหนึ่งเช่นกันว่า ค่าของความแปรปรวนของคะแนนในกลุ่มนักเรียนเหล่านี้จะค่อย ๆ เพิ่มมากขึ้นตามอายุด้วย. สาเหตุของเรื่องนี้คงเป็นทางองเดียวกันกับที่กล่าวมาแล้วในหน้า 38 ที่กล่าวว่า เด็กเล็ก ๆ จะมีความเป็นเอกพันธุ์ทั้งในด้านร่างกายและสมองมากกว่าเด็กโตที่ได้รับประสบการณ์แตกต่างกัน. ความแปรปรวนของคะแนนเหล่านี้จะมีมากที่สุดตรงระหว่างอายุ 11 ปี 7 เดือน ถึง 12 ปี 6 เดือน ซึ่งก็จะตรงกับเด็กนักเรียนระหว่างชั้นประถมปีที่ 4 และปีที่ 5 สภาพความจริงนี้สามารถอธิบายได้ง่าย เพราะเด็กในระดับอายุนี้ ประกอบด้วยนักเรียนประถมทั้งสองชั้นดังกล่าว คาบเรียนระหว่างเด็กที่อยู่ในเกณฑ์บังคับและนอกเกณฑ์บังคับของการศึกษา. ถ้าเป็นเด็กนอกเกณฑ์บังคับคือตรงตามประถมปีที่ 5 หรือสูงกว่าก็จะเป็นเด็กที่มีสภาพเป็นกลุ่มคัดสรร (Selected group) ที่มีสถานะ ทั้งทางการเงินและทางสติปัญญาสูงกว่าเด็กธรรมดาสามัญ เพราะโรงเรียนทั่วไปใช้วิธีสอบคัดเลือกเข้าเป็นนักเรียนชั้นประถมปีที่ 5. ฉะนั้น เด็กกลุ่มนี้จึงมีแนวโน้มของความสามารถแตกต่างกันมาก.

3. ถ้าจะประเมินคุณภาพของแบบทดสอบนี้ตามระดับอายุ ก็จะเห็นว่าเหมาะกับนักเรียนอายุ 11 ปี 7 เดือนถึง 12 ปี 6 เดือนมากที่สุด หรือก็คือนักเรียนประถมปีที่ 4 นั้นเอง เพราะเป็นกลุ่มที่มีความสามารถให้คะแนนเฉลี่ยใกล้เคียงกับครึ่งหนึ่งของคะแนนเต็มมากที่สุด คือให้คะแนนเฉลี่ย 19.93 คะแนน จากคะแนนเต็ม 40

4. สำหรับคุณภาพในทางความเชื่อมั่น (reliability) ของคะแนนนั้นก็ปรากฏว่า มีการแปรผันในทางเพิ่มขึ้นตามระดับอายุเช่นกัน. แบบทดสอบอุปมาอุปไมย-ตัวหนังสือนี้ สามารถให้ความเชื่อมั่นสูงสุดกับนักเรียนระดับอายุ 11 ปี 7 เดือน ถึง 12 ปี 6 เดือน ซึ่งมีความเชื่อมั่นถึง .853 ปรากฏการณ์สอดคล้องกับทฤษฎีทางการวัดผลเป็นอย่างดีว่า แบบทดสอบที่มีความแปรปรวนสูง ก็มักจะให้ความเชื่อมั่นสูงตามไปด้วย

ส่วนความคลาดเคลื่อนมาตรฐานของผลการทดสอบนั้น ก็เป็นที่น่ายินดีว่าทุกระดับอายุ จะมีความคลาดเคลื่อนของคะแนนคงที่อยู่ประมาณ ± 3 เสมอ ซึ่งก็หมายความว่า ประมาณ 68 % หรือ 2 ใน 3 ของเด็กที่ทำการทดลองนี้ จะมีคะแนนจริง (true score) ผิดแคบไปจาก คะแนนดิบ (observed score) ที่เขาสอบได้ ไม่เกิน ± 3 คะแนน.

5. คอยผลวิเคราะห์ต่าง ๆ ดังกล่าวมาข้างต้น เราจึงอาจประเมินคุณค่าแบบทดสอบฉบับนี้ โดยสรุปได้ว่า มีคุณภาพเหมาะสมสามารถจำแนกเด็กตามรายอายุได้ดีปาน ๆ กันกับที่จำแนกเด็ก เป็นรายชั้น.

ก. 6 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ จำแนกตามระดับอายุ

เพื่อทดสอบว่าความแปรผันของคะแนนเฉลี่ยของแต่ละระดับอายุทั้ง 9 ชั้น ดังที่ปรากฏใน ตารางที่ 11. นั้น จะมีความแตกต่างกันจริง ๆ อย่างมีนัยสำคัญเชิงสถิติหรือไม่ จึงได้นำตัวเลข เหล่านั้นมาวิเคราะห์ความแปรปรวนด้วยวิธี Analysis of Variance และทดสอบความแตกต่าง ซึ่งกันและกันด้วยวิธี t - test ดังปรากฏผลในตาราง 12. และ 13. ดังนี้

ตาราง 12. ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบอุปมาอุปไมย ประเภทตัวหนังสือ จำแนกตามรายอายุ

Source of Variation	Sum of Squares	df	Mean Square	F	p
Between Ages	25566.1288	8	3195.7661	90.6104	* * < .01
Within	46096.9981	1307	35.2693		
Total	71663.1269	1315			

ตาราง 13. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภท
ตัวหนังสือในแต่ละอายุ และความเชื่อมั่นของผลต่างเหล่านี้

ระดับอายุ	7	8	9	10	11	12	13	14	15
7	—	.5495	2.1111**	4.1213**	4.6354**	8.2002**	12.0570**	13.0666**	11.3565**
8		—	1.5615**	3.5713**	4.0859**	7.6507**	11.5075**	12.5191**	10.8070**
9			—	2.0102**	2.5243**	6.0891**	9.9459**	10.9555**	9.2454**
10				—	.5141	4.0789**	7.9357**	8.9453**	7.2352**
11					—	3.5648**	7.4216**	8.4312**	6.7211**
12						—	3.8568**	4.8664**	3.1563**
13							—	1.0096	.7005
14								—	1.7101
15									—

* แยกต่างอย่างมีนัยสำคัญที่ระดับน้อยกว่า .05

** แยกต่างอย่างมีนัยสำคัญที่ระดับน้อยกว่า .01

จากตารางที่ 12. ปรากฏว่า ถ้าวิเคราะห์โดยส่วนรวมแล้วความแปรปรวนของคะแนนเฉลี่ยในแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือมีความแตกต่างอย่างมีนัยสำคัญเชิงสถิติอย่างเชื่อมั่น จึงได้ทำการทดสอบความแตกต่างของคะแนนเฉลี่ยระหว่างอายุ ด้วยวิธี $t - test$ เพื่อตรวจสอบว่าคะแนนเฉลี่ยของอายุคู่ใดจะมีความแตกต่างกันอย่างเชื่อมั่นได้บ้าง ซึ่งก็ปรากฏว่า ส่วนใหญ่ของคะแนนเฉลี่ยเหล่านี้มีความมากกว่ากันจนเป็นเหตุเชื่อถือได้ ยกเว้นระหว่างอายุ 7 กับ 8 ปี, 10 กับ 11 ปี, 13 กับ 14 ปี, 13 กับ 15 ปี, และ 14 กับ 15 ปี เพียง 5 คู่ เท่านั้นที่แตกต่างกันเพียงเล็กน้อย ยังไม่ถึงขนาดที่เชื่อมั่นได้. ซึ่งก็หมายความว่า เด็กระดับเหล่านี้คงมีความสามารถในการทำแบบทดสอบอุปมาอุปไมยตัวหนังสือปาน ๆ กันยังไม่เคยค่อยแตกต่างกันจริง

การที่จะแนบเฉลี่ยสูงต่ำกว่ากันนั้นคงเป็นไปตามโอกาส เนื่องจากความกระเพื่อม (fluctuation) ของกลุ่มตัวอย่างมากกว่า. เป็นที่น่าสังเกตว่า ความสามารถในการคำนวณโดยไม่ตัวหนังสือซึ่งเป็นส่วนหนึ่งของสมรรถภาพด้านวิจารณ์ (Reasoning) ในกลุ่มเด็กตัวแบบนี้จะต้องใช้เวลาพัฒนาประมาณ 2 ปี จึงจะมีขีดความสามารถแตกต่างจากเดิมอย่างมีนัยได้.

ก. 7 คุณภาพของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกเป็นรายอายุ

การวิเคราะห์คุณภาพของแบบทดสอบฉบับนี้ดำเนินการแบบเดียวกับที่ได้กระทำมาแล้วในข้อก่อน คือนำข้อมูลมาแจกเป็น 9 กลุ่มแล้ววิเคราะห์หาค่าสถิติพื้นฐานของแต่ละกลุ่ม เพื่อหาความเหมาะสมและความเชื่อมั่น ดังตารางข้างล่างนี้

ตาราง 14. ค่าสถิติพื้นฐานของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกเป็นรายอายุ

ระดับอายุเป็นปี	N	$\bar{X}$	T-score	s^2	s	r_{tt}	$SE_{meas.}$
7	74	14.4865	42	17.4861	4.1816	.484	± 3.0049
8	154	14.7273	43	10.9316	3.3063	.153	± 3.0428
9	157	16.8089	46	32.8735	5.7335	.722	± 3.0256
10	208	18.8269	49	33.7863	5.8126	.723	± 3.0586
11	189	19.0638	50	24.9260	4.9926	.615	± 3.0974
12	157	22.4522	53	45.2108	6.7239	.802	± 2.9908
13	211	25.3555	56	37.1064	6.0915	.769	± 2.9276
14	108	25.8889	57	41.2399	7.4218	.799	± 3.3309
15	58	25.8621	57	29.9456	5.4723	.713	± 2.9337

ตารางนี้ ได้ให้ความหมายที่สำคัญ 5 ประการคือ

1. คะแนนเฉลี่ยของนักเรียนแต่ละระดับอายุมีแนวโน้มที่จะเพิ่มสูงขึ้นเรื่อย ๆ ตามระดับอายุของนักเรียนแบบเดียวกันกับแบบทดสอบอุปมาอุปไมย-ประเภทตัวหนังสือ คือกลุ่มนักเรียนที่มีอายุมาก จะมีความสามารถในการทำแบบทดสอบนี้ได้ดีกว่ากลุ่มที่มีอายุน้อย. คุณลักษณะนี้เป็นเครื่องยืนยันถึงความไวของแบบทดสอบในการจำแนกสมรรถภาพของนักเรียนแต่ละระดับได้.

ถ้าพิจารณาจากคะแนนเฉลี่ยในรูปของคะแนน T ในแต่ละระดับอายุก็จะเห็นว่า อัตราการเพิ่มของคะแนน T นี้เป็นไปในทำนองเดียวกันกับแบบทดสอบฉบับแรก คือจะค่อย ๆ ทวีจาก 1 หน่วยในระดับอายุ 7 และ 8 ขวบ จนถึง 3 หน่วยในระดับอายุ 12 และ 13 ขวบ.

2. ค่าสถิติที่สำคัญอีกชนิดหนึ่งก็คือ ค่าความแปรปรวนของคะแนนในกลุ่มนักเรียนเหล่านั้น ซึ่งจะค่อย ๆ เพิ่มขึ้นตามอายุแบบเดียวกันกับคะแนนเฉลี่ย. สาเหตุของเรื่องนี้เนื่องจากแบบทดสอบอุปมาอุปไมยทั้ง 2 ฉบับมีเนื้อหาและอื่น ๆ คล้ายกันจนอาจจะกล่าวได้ว่าเป็นแบบทดสอบคู่ขนาน (Parallel form) กัน. ฉะนั้น ค่าความแปรปรวนของแบบทดสอบจึงเป็นไปตามทำนองเดียวกัน คือจะแปรปรวนสูงสุดตรงระดับเด็กอายุ 11 ปี 7 เดือน ถึง 12 ปี 6 เดือน ซึ่งเด็กในระดับอายุส่วนใหญ่เป็นเด็กที่อยู่ในชั้นประถมศึกษาปีที่ 5.

3. คุณค่าของแบบทดสอบนี้จะเหมาะสมที่สุดกับกลุ่มเด็กอายุ 12 ปี เพราะเป็นกลุ่มที่ได้คะแนนเฉลี่ย 19.07 ใกล้เคียงกับครึ่งหนึ่งของคะแนนเต็มมากที่สุด.

4. สำหรับคุณภาพของแบบทดสอบในด้านการเชื่อมั่นของคะแนนนั้นจะเห็นว่า มีการเปลี่ยนแปลงไปตามระดับอายุเช่นกัน ซึ่งระดับอายุที่สามารถให้ค่าความเชื่อมั่นสูงสุดคือ 11 ปี 7 เดือน ถึง 12 ปี 6 เดือน เช่นเดียวกับแบบทดสอบอุปมาอุปไมย-ประเภทตัวหนังสือ. ความเชื่อมั่นนั้นค่าถึง .802 ซึ่งนับว่าเป็นความเชื่อมั่นที่ค่อนข้างสูง. ส่วนความคลาดเคลื่อนมาตรฐานของแบบทดสอบนี้ก็มีค่าประมาณ ± 3.33 โดยตลอดทุกอายุ สูงกว่าฉบับก่อนเล็กน้อย.

5. จากผลวิเคราะห์ดังกล่าวข้างต้นจึงอาจสรุปได้ว่า แบบทดสอบฉบับนี้มีคุณภาพดีพอ ๆ กับแบบทดสอบอุปมาอุปไมยประเภทภาษา.

ก. 8 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

จำแนกตามระดับอายุ

จากผลการวิเคราะห์ในข้อ ก. 7 แสดงว่า ในแต่ละระดับอายุจะได้คะแนนเฉลี่ยมากน้อยต่าง ๆ กัน

จากนี้จึงทำการวิเคราะห์หาความแตกต่างของคะแนนเฉลี่ยและทดสอบนัยสำคัญของความแตกต่างนั้นด้วยวิธี t -test แบบเดียวกับที่เคยกระทำมาแล้วในแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ ซึ่งปรากฏผลดังแสดงไว้ในตารางที่ 15. และที่ 16. ดังนี้

ตาราง 15. ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ จำแนกตามวัยอายุ

Source of Variation	Sum of Squares	df	Mean Square	F	p
Between	21220.9437	8	2652.6179	85.1377	< .01
Within	40721.9712	1307	31.1568		
Total	61942.9149	1315			

ตาราง 16. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบอุปมาอุปไมยรูปภาพ
ในแต่ละอายุและระดับความเชื่อมั่นของผลต่างเหล่านั้น

ระดับอายุ	7	8	9	10	11	12	13	14	15
7	—	.2408	1.8224**	4.3404**	4.5823**	7.9661**	10.8690**	11.4024**	11.3756**
8		—	2.0816**	4.0996**	4.3415**	7.7253**	10.6282**	11.1616**	11.1348**
9			—	2.0180**	2.2599**	5.6433**	8.5466**	9.0800**	9.0532**
10				—	.2419	3.6257**	6.5286**	7.0620**	7.0352**
11					—	3.3838**	6.2867**	6.8401**	6.7933**
12						—	2.9029**	3.4363**	3.4095**
13							—	.5334	.5066
14								—	.0268
15									—

จากตาราง 15. ปรากฏว่า F มีค่ามากถึง 85.1377 จึงแปลความหมายได้ว่า คะแนนเฉลี่ยมีความแตกต่างจากกันอย่างมีนัยสำคัญเชิงสถิติ จึงสมควรที่จะทดสอบให้แน่ชัดว่าคะแนนเฉลี่ยระหว่างอายุคู่ใดมีความแตกต่างกันจริง จนเชื่อถือได้ดังที่แสดงไว้ในตารางที่ 16. นั้น ผลการเปรียบเทียบเป็นราย ๆ คู่ ปรากฏว่า ส่วนใหญ่ของคะแนนเฉลี่ยในแต่ละระดับอายุแตกต่างกัน นอกจากอายุ 2 ระดับ อายุที่ติดต่อกันคือ อายุ 7 กับ 8 ขวบ, 10 กับ 11 ขวบ, 13 กับ 14 ขวบ, 13 กับ 15 ขวบ, และ 14 กับ 15 ขวบ รวม 5 คู่. ผลการทดสอบให้ผลเป็นทำนองเดียวกับแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ.

ก. 9 คุณภาพของแบบทดสอบไม่เข้าพวกที่เป็นตัวหนังสือ จำแนกเป็นรายอายุ

คุณภาพของแบบทดสอบฉบับนี้ใกล้เคียงกับ 2 ฉบับที่กล่าวมาแล้ว คือนำข้อมูลมาจำแนกเป็นรายอายุ แล้วทำการวิเคราะห์หาค่าสถิติพื้นฐานของแบบทดสอบนั้น ดังตารางที่ 17. ข้างล่างนี้

ตาราง 17. คาสถิติพื้นฐานของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ จำแนก
เป็นรายอายุ

ระดับอายุเป็นปี	N	$\bar{X}$	T-score	s^2	s	r_{tt}	$SE_{meas.}$
7	74	11.4324	41	17.8378	4.2235	.556	± 2.8137
8	154	12.2013	43	18.0441	4.2478	.544	± 2.8690
9	157	14.6115	47	32.2390	5.6779	.731	± 2.9468
10	208	16.4760	48	34.3085	5.8573	.736	± 3.0101
11	189	16.4815	49	33.8892	5.8214	.732	± 3.0114
12	157	20.4140	53	51.8980	7.2040	.828	± 2.9868
13	211	23.5735	56	34.9029	5.9079	.741	± 3.0018
14	108	23.9815	57	29.3454	5.4171	.689	± 3.0162
15	58	24.1552	58	35.2211	5.9347	.747	± 2.9852

การวิเคราะห์ข้อมูลนี้ปรากฏว่า ได้ผลยืนยันเป็นทำนองเดียวกับของแบบทดสอบฉบับมาตรฐาน
กล่าวคือ

1. คะแนนเฉลี่ยจะมีค่าเพิ่มสูงขึ้นตามลำดับอายุที่มากขึ้น และแบบทดสอบนี้มีความง่าย
เหมาะกับกลุ่มนักเรียนระดับอายุ 12 ปีมากที่สุด ถ้าใช้กับกลุ่มอายุน้อยกว่านี้ จะเป็นแบบทดสอบ
ที่ยากเกินไป และกับกลุ่มอายุที่มากกว่านี้ก็จะกลายเป็นแบบทดสอบง่ายเกินไป.

2. ค่าความแปรปรวนของคะแนนในแต่ละกลุ่มอายุก็จะเพิ่มสูงขึ้นเรื่อย ๆ จากอายุ 7 ปี
ที่มีค่าเพียง 17.8378 จนถึงระดับสูงสุดที่อายุ 12 ปี ซึ่งมีค่า 51.8980 ถ้าอายุมากกว่านั้น
ค่าความแปรปรวนจะค่อย ๆ ลดลง อันแสดงว่าความสามารถในการจัดเข้าพวกของกลุ่มนักเรียน
ที่มีอายุมาก จะเป็นเอกพันธ์กันมากยิ่งขึ้น.

3. ความเชื่อมั่นของแบบทดสอบนี้ในแต่ละระดับอายุ ก็ปรากฏว่า มีความเปลี่ยนแปลงจากน้อยไปมากที่สุดที่อายุ 12 ขวบ แล้วมีแนวโน้มค่อย ๆ ลดลงเป็นทำนองเดียวกันกับความแปรปรวนเช่นกัน. ส่วนความคลาดเคลื่อนมาตรฐานของแบบทดสอบนี้ มีเพียงประมาณ 3 คะแนน ขนาดเกี่ยวข้องกับอุปมาอุปไมยประเภทตัวหนังสือ

ก. 10 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกตัวหนังสือ จำแนกตามระดับอายุ

จากการวิเคราะห์เบื้องต้น ก็ได้นำข้อมูลเหล่านั้นมาเปรียบเทียบความแตกต่างของคะแนนเฉลี่ยในแต่ละระดับอายุ และทดสอบความแตกต่างนั้นด้วย t - test ดังตารางที่ 18. และที่ 19.

ตาราง 18. ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ จำแนกตามรายอายุ

Source of Variation	Sum of Squares	df	Mean Square	F	p
Between	24521.0414	8	3065.1301	92.6519	< .01
Within	43238.5536	1307	33.0822		
Total	67759.5950	1315			

ตาราง 19. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกเป็นตัวหนังสือ
ในแต่ละอายุและความเชื่อมั่นของผลต่างเหล่านั้น

ระดับอายุ	7	8	9	10	11	12	13	14	15
7	—	.7689	3.1691**	5.0436**	5.0491**	8.9897**	12.1411**	12.5491**	12.7228**
8	—	—	2.4102**	4.2753**	4.2802**	8.2127**	11.3722**	11.7802**	11.9539**
9	—	—	—	1.8645**	1.8700**	5.8025**	8.9620**	9.3700**	9.5437**
10	—	—	—	—	.0055	3.9380**	7.0975**	7.5055**	7.6790**
11	—	—	—	—	—	3.9325**	7.0920**	7.5000**	7.6737**
12	—	—	—	—	—	—	3.1595**	3.5675**	3.7412**
13	—	—	—	—	—	—	—	.4080	.5817
14	—	—	—	—	—	—	—	—	.1737
15	—	—	—	—	—	—	—	—	—

* แตกต่างอย่างมีนัยสำคัญที่ระดับน้อยกว่า .05

** แตกต่างอย่างมีนัยสำคัญที่ระดับน้อยกว่า .01

ผลการวิเคราะห์และผลการทดสอบ ปรากฏว่า ส่วนใหญ่ของคะแนนเฉลี่ยของกลุ่มอายุเหล่านี้จะแตกต่างอย่างมีนัยสำคัญเชิงสถิติจนเป็นที่เชื่อถือได้ นอกจากระหว่างอายุ 7 กับ 8 ขวบ, 10 กับ 11 ขวบ, 13 กับ 14 ขวบ, 13 กับ 15 ขวบ, และ 14 กับ 15 ขวบ เพียง 5 คู่เท่านั้นที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญที่เชื่อมั่นได้

ก. 11 คุณภาพของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกตามรายอายุ

การวิจัยนี้ได้นำข้อมูลจากแบบทดสอบไม่เข้าพวกที่เป็นรูปภาพมาวิเคราะห์เพื่อหาความสัมพันธ์ต่าง ๆ ดังที่ปฏิบัติมาแล้วกับ 3 ฉบับก่อน เริ่มด้วยจำแนกข้อมูลเป็นรายอายุ แล้วหาคะแนนเฉลี่ยของแต่ละกลุ่มอายุนั้น ๆ พร้อมกับหาความแปรปรวน, ความเชื่อมั่นของแบบทดสอบ, และความคลาดเคลื่อนมาตรฐานของคะแนนสอบด้วย ดังที่แสดงไว้ในตารางที่ 20.

ตาราง 20. การศึกษาพื้นฐานของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ จำแนกเป็นรายอายุ

ระดับอายุเป็นปี	N	$\bar{X}$	T-score	s^2	s	r_{tt}	$SE_{meas.}$
7	74	16.6892	40	30.3815	5.5119	.697	± 3.0321
8	154	18.3571	42	22.3748	4.7302	.570	± 3.1002
9	157	20.2994	45	32.2751	5.6811	.708	± 3.0701
10	208	22.4231	48	27.7525	5.2681	.662	± 3.0650
11	189	22.5979	49	24.4331	4.9430	.613	± 3.1146
12	157	24.9554	52	30.8249	5.5520	.713	± 2.9725
13	211	26.3697	54	16.5960	4.0738	.470	± 2.9649
14	108	26.7130	56	23.0851	4.8047	.631	± 2.9169
15	58	27.2241	57	24.7384	4.9738	.665	± 2.8783

ผลการวิเคราะห์พอจะสรุปความสำคัญได้ ดังนี้

1. คะแนนเฉลี่ยของแบบทดสอบนี้เพิ่มขึ้นตามระดับอายุแบบเดียวกับแบบทดสอบอื่น ๆ และกลุ่มอายุที่มีความสามารถเหมาะสมกับแบบทดสอบฉบับนี้คืออายุ 9 ขวบ ซึ่งต่างกับแบบทดสอบไม่เข้าพวกที่เป็นตัวหนังสือและอุปมาอุปไมย. ผลวิเคราะห์หนึ่งให้แน่ชัดว่าแบบทดสอบไม่เข้าพวกประเภทรูปภาพนี้มีระดับความยากน้อยกว่า แบบทดสอบฉบับอื่น ๆ หรืออีกนัยหนึ่งนี้อาจพูดกลับได้ว่า แบบทดสอบที่เป็นรูปภาพนี้เหมาะกับเด็กเล็ก ๆ มากกว่าแบบทดสอบที่เป็นตัวหนังสือก็ได้ ทั้งนี้เพราะทั้งแบบทดสอบที่เป็นตัวหนังสือและรูปภาพนี้ต่างก็วัดเนื้อหาเรื่องเดียวกัน.

2. ส่วนค่าความแปรปรวน (Variance) ของคะแนนในแบบทดสอบนั้นปรากฏว่า ไม่เพิ่มขึ้นหรือลดลงตามลำดับเรียงแบบทดสอบชนิดใด ๆ บางกลุ่มอายุอาจแปรปรวนมากหรือน้อยกว่าอายุที่ติด

ไปได้. ความไม่คงที่นี้ทำให้เราทราบว่า ในแต่ละกลุ่มอายุมีเด็กที่มีความสามารถในการ
จัดเซาพวกมากน้อยแตกต่างกัน.

3. ในด้านความเชื่อมั่นของแบบทดสอบฉบับนี้ในแต่ละระดับอายุ ปรากฏว่าจะมีค่าความ
เชื่อมั่นระหว่าง .470 - .708 และในการนี้กลุ่มระดับอายุที่มีความเชื่อมั่นสูงสุดคืออายุ 12 ปี
ตรงกับแบบทดสอบฉบับอื่น ๆ สำหรับความคลาดเคลื่อนมาตรฐานของแบบทดสอบนี้ ก็จะอยู่ในระหว่าง
 ± 3 คะแนน เช่นเดียวกันกับแบบทดสอบอื่น ๆ.

ก. 12 การทดสอบความแตกต่างของคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกที่เป็นรูปภาพ จำแนก
ตามรายอายุ

จากนี้ ก็ได้ทำการวิเคราะห์หาความแปรปรวนของคะแนนเฉลี่ย และทดสอบนัยสำคัญของ
ความแตกต่างด้วย $t - test$ ดังแสดงไว้ในตารางที่ 21. และที่ 22.

ตาราง 21. ผลการวิเคราะห์ความแปรปรวนของคะแนนจากแบบทดสอบไม่เข้าพวก
รูปภาพ จำแนกรายอายุ

Source of Variation	Sum of Squares	df	Mean Square	F	p
Between	13017.1398	8	1627.1424	64.0791	< .01
Within	33188.3891	1307	25.3927		
Total	46205.5289	1315			

ตาราง 22. ความแตกต่างระหว่างคะแนนเฉลี่ยของแบบทดสอบไม่เข้าพวกรูปภาพ
ในแต่ละอายุและความเชื่อมั่นของผลต่างเหล่านั้น

ระดับอายุ	7	8	9	10	11	12	13	14	15
7	—	1.6679*	3.6102**	5.7339**	5.9087**	8.2662**	9.6805**	10.0238**	10.5349**
8		—	1.9423**	4.0660**	4.2408**	6.5983**	8.0126**	8.3559**	8.8670**
9			—	2.1237**	2.2985**	4.6560**	6.0703**	6.4136**	6.9247**
10				—	.1748	2.5323**	3.9466**	4.2899**	4.8010**
11					—	2.3575**	3.7718**	4.1151**	4.6262**
12						—	1.4143**	1.7576**	2.2687**
13							—	.3433	.8544
14								—	.5111
15									—

ผลการวิเคราะห์ความแปรปรวนปรากฏว่า คะแนนเฉลี่ยโดยส่วนรวมทั้ง 9 ระดับอายุนี้ มีความแตกต่างกันอย่างมีนัยสำคัญ และเมื่อทดสอบความแตกต่างกันด้วย t -test แล้ว ก็ปรากฏว่า ส่วนใหญ่ของคะแนนเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญเชิงสถิติ คือนักเรียนแต่ละกลุ่มอายุ มีความสามารถในการจำเข้าพวกสูงต่ำกว่ากันจริงจนเป็นที่เชื่อมั่นได้ ยกเว้นความสามารถระหว่าง ระดับอายุ 10 กับ 11 ขวบ, 13 กับ 14 ขวบ, 14 กับ 15 ขวบ, เพียง 3 คู่เท่านั้นที่ไม่ แตกต่างกันอย่างมีนัยสำคัญ.

ก. 13 คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 2 จำแนกตามเพศ

การวิเคราะห์ที่แล้มา เป็นการเปรียบเทียบคะแนนเฉลี่ยโดยปีชั้นและอายุเป็นเกณฑ์ การศึกษาที่จะกระทำต่อไปนี้ เป็นการวิเคราะห์โดยยึดเพศเป็นเกณฑ์ ซึ่งในการนี้ได้อำนาจโดย นำข้อมูลทั้งหมด 1316 คน มาแยกเป็น 2 กลุ่มย่อยตามเพศ แล้วคำนวณค่าเชิงสถิติของแต่ละกลุ่ม เหล่านั้นเป็นรายชั้นเรียน. ดังจะได้เสนอผลการวิจัยของกลุ่มนักเรียนชั้นประถมศึกษาปีที่ 2 ดังตาราง 23.

ตาราง 23. ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับ ของนักเรียนชั้นประถมศึกษาปีที่ 2
จำแนกตามเพศ

ค่าสถิติตามเพศ	อุปมาอุปไมย ตัวหนังสือ	อุปมาอุปไมย ภาพ	ไม่เขาพวก ตัวหนังสือ	ไม่เขาพวก ภาพ
จำนวนนักเรียนชาย	232	232	232	232
$\bar{X}$	12.1595 = T 43	14.5647 = T42	12.3319 = T43	18.5302 = T42
s^2	18.7926	17.4849	17.2703	26.9802
r_{tt}	.5637	.4823	.5190	.6475
$SE_{meas.}$	± 2.8633	± 2.9919	± 2.8820	± 3.0838
จำนวนนักเรียนหญิง	192	192	192	192
$\bar{X}$	11.5417 = T 42	14.2448 = T41	11.4479 = T42	17.8281 = T41
s^2	14.0820	12.5732	15.6307	24.8341
r_{tt}	.4275	.2774	.4894	.6175
$SE_{meas.}$	± 2.8392	± 3.0144	± 2.8248	± 3.0817
ผลต่าง $\bar{X}$	.6178	.3199	.8840	.7229
$SE_{diff.}$	.3928	.3753	.3947	.5080
t	1.573	.852	2.240 *	2.423 *
F	1.34 *	1.39 **	1.11	1.09

ตารางนี้ให้ความหมายเชิงเปรียบเทียบระหว่างนักเรียนชาย - หญิงของชั้นประถมศึกษาปีที่ 2
ไต่หลายประการ ดังนี้

- คะแนนเฉลี่ยในแบบทดสอบทุกฉบับของนักเรียนชาย สูงกว่าของนักเรียนหญิงเสมอ.
และถ้าใช้คะแนน T ปรกติเป็นหลักแห่งการวัดแล้วคะแนน T ของนักเรียนชายจะมากกว่าของหญิง

อยู่ประมาณ 1 หน่วยทุกฉบับ แต่อย่างไรก็ดี เมื่อนำค่าแตกต่างของคะแนนเหล่านี้มา ทดสอบด้วย t - test แล้ว ปรากฏว่า มีแต่เพียงแบบทดสอบไม่เข้าพวกทั้งประเภทตัวหนังสือและรูปภาพ 2 ฉบับเท่านั้น ที่นักเรียนชายทำคะแนนได้สูงกว่าหรือมีความสามารถเหนือกว่านักเรียนหญิง อย่างเชื่อมั่นได้ ณ ระดับ 5 %

2. ปริมาณความแปรปรวนของคะแนนในกลุ่มนักเรียนชายจะมากกว่าของนักเรียนหญิงเสมอ ซึ่งแสดงว่า กลุ่มนักเรียนชายประกอบไปด้วยเด็กที่มีความรู้ทั้งที่เก่งมากและอ่อนมากปนคละกัน ส่วนความสามารถของเด็กนักเรียนหญิงจะสม่ำเสมอใกล้เคียงกันมากกว่านักเรียนชาย แต่อย่างไรก็ดี เมื่อนำความแปรปรวนระหว่างนักเรียนทั้งสองเพศมา เปรียบเทียบกันแล้ว ก็ปรากฏว่ามีเพียงคะแนนจากแบบทดสอบอุปมาอุปไมยทั้งประเภทตัวหนังสือและรูปภาพ เท่านั้น ที่แต่ละเพศมีความแปรปรวนแตกต่างกันอย่างมีนัยสำคัญเชิงสถิติ ส่วนความแปรปรวนในแบบทดสอบไม่เข้าพวกทั้ง 2 ฉบับนั้น ไม่แตกต่างจากกันจนเห็นชัดนัก ซึ่งก็หมายความว่า ทั้งนักเรียนชายและหญิงต่างมีอัตราการกระจายของคะแนนในแบบทดสอบไม่เข้าพวกเป็นปริมาณปาน ๆ กัน.

3. ในด้านความเชื่อมั่นของคะแนนสอบก็เป็นที่น่าสนใจ เพราะว่าในกลุ่มนักเรียนชายจะมีความเชื่อมั่นได้สูงกว่านักเรียนหญิงทุกฉบับ ทั้ง ๆ ที่คะแนนความคลาดเคลื่อนมาตรฐานเกือบจะเท่ากันก็ตาม.

ก. 14 คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 4 จำแนกตามเพศ

โดยทำนองเดียวกันกับข้อก่อน ให้นำข้อมูลของนักเรียนชั้นประถมศึกษาปีที่ 4 มาแยกตามเพศ แล้วคำนวณหาค่าสถิติพื้นฐาน ดังแสดงไว้ในตารางที่ 24. ดังนี้

ตาราง 24. คาสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับของนักเรียนชั้นประถมศึกษาปีที่ 4
จำแนกตามเพศ

คาสถิติตามเพศ	อุปมาอุปไมย ตัวหนังสือ	อุปมาอุปไมย ภาพ	ไม่เขาพวก ตัวหนังสือ	ไม่เขาพวก ภาพ
จำนวน ชาย	259	259	259	259
$\bar{X}$	17.2857 = T 50	20.3161 = T 51	18.0502 = T 51	27.7967 = T 58
s^2	36.6777	26.2094	29.7686	17.3381
r_{tt}	.7175	.6344	.6843	.5239
$SE_{meas.}$	± 3.0383	± 3.0952	± 3.0652	± 3.1151
จำนวน หญิง	208	208	208	208
$\bar{X}$	16.7516 = T 49	19.1731 = T 49	16.5096 = T 49	26.5307 = T 55
s^2	32.4145	25.9119	34.2801	14.7223
r_{tt}	.7175	.6304	.7355	.4032
$SE_{meas.}$	± 2.9803	± 3.0445	± 3.0106	± 3.9743
ผลต่าง $\bar{X}$	1.1992	1.1435	1.5405	1.2660
$SE_{diff.}$	.3573	.4751	.5288	.3907
t	3.356**	2.407*	2.913**	3.240**
F	1.13	1.01	1.15	1.18

การวิเคราะห์เนื้อหาขอเท็จจริงที่สำคัญหลายประการ เช่นเดียวกับตารางก่อน ๆ คือ

1. นักเรียนชั้นประถมศึกษาปีที่ 4 ชาย โดยส่วนรวม จะมีความสามารถเหนือกว่านักเรียนหญิงในแบบทดสอบทุกฉบับ อย่างมั่นใจได้
2. ความแปรปรวนของคะแนนสอบ ของนักเรียนทั้ง 2 เพศนี้ ไม่แตกต่างจากกันจนถึงขั้นมีนัยสำคัญเชิงสถิติเลย ซึ่งก็หมายความว่านักเรียน 2 กลุ่มนี้มีอัตราการกระจายของคะแนนเป็น

แบบเดียวกัน จนจัดว่า เป็นเอกพันธุ์ได้.

3. เป็นที่น่าสังเกตว่า ความเชื่อมั่นของแบบทดสอบอุปมาอุปไมยของนักเรียนชายและหญิงต่างจะมีค่าเกือบเท่ากัน ทั้ง ๆ ที่อุปมาอุปไมยประเภทภาพ จะมีค่าน้อยกว่าประเภทตัวหนังสืออยู่บ้างก็ตาม และความเชื่อมั่นในแบบทดสอบไมใช่ภาพประเภทภาพ จะมีค่าน้อยกว่าแบบอื่น ๆ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4.

ก. 15 คุณภาพของแบบทดสอบในนักเรียนชั้นประถมศึกษาปีที่ 7 จำแนกตามเพศ

ในการวิเคราะห์ขั้นสุดท้ายเกี่ยวกับอิทธิพลของเพศที่อาจมีต่อคุณภาพของแบบทดสอบทั้ง 4 ฉบับนั้น ได้กระทำโดยแจกแจงข้อมูลของนักเรียนชั้นประถมศึกษาปีที่ 7 ออกเป็น 2 กลุ่มตามเพศ พร้อมกับหาค่าสถิติดังแสดงในตารางที่ 25.



ตาราง 25. ค่าสถิติพื้นฐานของแบบทดสอบทั้ง 4 ฉบับของนักเรียนชั้นประถมศึกษาปีที่ 7

จำแนกตามเพศ

ค่าสถิติตามเพศ	อุปมาอุปไมย ตัวหนังสือ	อุปมาอุปไมย ภาพ	ไม่เข้าพวก ตัวหนังสือ	ไม่เข้าพวก ภาพ
จำนวน ชาย	246	246	246	246
$\bar{X}$	25.1220 = T 58	27.1423 = T 58	25.2683 = T 58	27.7967 = T 57
s^2	30.7442	27.6817	23.4624	17.3381
s	.7139	.7023	.6188	.5239
SE _{meas.}	± 2.9653	± 2.8706	± 2.9906	± 2.8731
จำนวน หญิง	179	179	179	179
$\bar{X}$	24.5642 = T 57	25.8603 = T 57	24.2849 = T 57	26.5307 = T 55
s^2	29.2247	25.8961	22.6767	14.7203
s	.6929	.6635	.5941	.4032
SE _{meas.}	± 2.9953	± 2.9515	± 3.0339	± 2.9640
ผลต่าง $\bar{X}$	.5578	1.2820	.9834	1.2660
SE _{diff.}	.5368	.5071	.4712	.3907
t	1.039	2.528 *	2.087 *	3.240 **
F	1.05	1.07	1.03	1.18

จากตารางนี้ จะกล่าวถึงค่าสถิติที่แสดงคุณภาพของแบบทดสอบฉบับนี้ คือ

1. ทั้งนักเรียนชายและหญิงจะได้คะแนนเฉลี่ยแตกต่างกันทุกฉบับ โดยที่คะแนนเฉลี่ยของนักเรียนชายจะมากกว่านักเรียนหญิงเสมอ และเมื่อทดสอบนัยสำคัญเชิงสถิติแล้ว ปรากฏว่าทั้งแบบทดสอบอุปมาอุปไมยประเภทรูปภาพไม่เข้าพวกประเภทที่เป็นตัวหนังสือ และไม่เข้าพวก

ประเภทที่เป็นรูปภาพรวม 3 ฉบับจะมีคะแนนเฉลี่ยของนักเรียนหญิงและนักเรียนชายแตกต่างกันจริง ส่วนแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือจะมีคะแนนเฉลี่ยระหว่างนักเรียนชายและหญิงแตกต่างกันอย่างไม่มีนัยสำคัญ.

2. ค่าความแปรปรวนของคะแนน ทั้งนักเรียนชายและนักเรียนหญิงจะมีค่าความแปรปรวนไม่แตกต่างกันเลยทุกฉบับ นั้นแสดงว่า กลุ่มนักเรียนชายและหญิงจะมีความสามารถในการทำแบบทดสอบเหล่านี้มากน้อยปาน ๆ กัน.

3. ในด้านความเชื่อมั่นของแบบทดสอบ จะเห็นว่า กลุ่มนักเรียนชายจะมีความเชื่อมั่นของแบบทดสอบทุกฉบับมากกว่านักเรียนหญิง และค่าความเชื่อมั่นเหล่านี้จะมีค่าความคลาดเคลื่อนมาตรฐานเพียงประมาณ ± 3 คะแนนเท่านั้น ซึ่งอาจสรุปได้ว่า แบบทดสอบทั้ง 4 ฉบับมีคุณภาพดีพอที่จะจำแนกกลุ่มนักเรียนชายและหญิงได้.

ข. คุณภาพของแบบทดสอบในด้านอำนาจจำแนก

คุณภาพที่สำคัญยิ่งอีกประการหนึ่งของแบบทดสอบ นอกจากระดับความยากง่ายแล้ว ก็คือ อำนาจจำแนกของแบบทดสอบนั้น. คุณสมบัตินี้หมายถึงความสามารถของแบบทดสอบที่จะให้คะแนนหรือผลการสอบของเด็กเก่ง - เด็กอ่อน หรือของเด็กอายุมาก - อายุน้อย แตกต่างกันอย่างเด่นชัด.

ความยากง่าย (Difficulty) และอำนาจจำแนก (Discrimination) นี้ไม่เหมือนกัน ต่างมีความหมายและขอบเขตแตกต่างกัน. อย่างแรกหมายถึงปริมาณนักเรียนที่ตอบข้อนั้นถูกมากหรือน้อย ถ้าข้อใดเด็กทำข้อนั้นถูกมาก ก็แปลว่าข้อนั้นง่ายมาก หรือหมายความว่ามีความยากน้อย. และโดยทำนองเดียวกัน ถ้าจะกล่าวถึงระดับความยากง่ายของแบบทดสอบใดทั้งฉบับ ก็หมายความว่า เด็กนักเรียนกลุ่มนั้นสามารถทำคะแนนได้คะแนนสูงลิ้นจี่จากแบบทดสอบฉบับนั้นมากหรือน้อย มักได้คะแนนใกล้เคียงศูนย์หรือคะแนนจวนเต็ม เป็นต้น. วิธีหาค่าความยากง่ายของแบบทดสอบ ก็อาจจะทำได้โดยคิดสัดส่วนของจำนวนผู้สอบที่ทำคำถามข้อใดถูก หรือคิดความยากเป็นหน่วยของคะแนนมาตรฐานทางเหนือจากคะแนนเฉลี่ยไปเท่าใด เป็นต้น. ถ้าเป็นการหาระดับความยากของแบบทดสอบทั้งฉบับ ก็อาจหาได้โดยวิธีคำนวณหารายเฉลี่ยของความยากมาตรฐาน (Δ) ของแต่ละข้อคำถามทั้งฉบับ หรือใช้วิธีตรวจสอบดูว่าคะแนนเฉลี่ยของแบบทดสอบนั้น ห่างจากครึ่งหนึ่งของคะแนนเต็ม (possible score) ไม่มากน้อยเพียงใด ซึ่งเป็นวิธีที่ใช้ในรายงานฉบับนี้.

ส่วนอำนาจจำแนกหมายถึงความสัมพันธ์สอดคล้องระหว่างเด็กเก่งที่คะแนนสูง กับที่เด็กทำข้อสอบหรือผิด ตามผลการวิเคราะห์ปรากฏว่า เด็กเก่งมักทำคำถามข้อใดถูกมาก พร้อม ๆ กับที่เด็กอ่อนก็มักจะตอบข้อนั้นผิดมากด้วยแล้ว คำถามข้อนั้นก็อำนาจจำแนกสูง และมีค่าเป็นบวก ถ้าทั้งเด็กเก่งและเด็กอ่อนต่างก็ตอบข้อใดถูกหรือผิดเป็นจำนวนเท่า ๆ กัน คำถามข้อนั้นก็ไม่มีอำนาจจำแนกเลย คือไม่สามารถบอกให้เรารู้ว่า ผู้ที่ตอบถูกก็หรือผู้ที่ตอบผิดก็จะเป็นเด็กเก่งหรืออ่อนแน่ และโดยตรงข้าม ถ้าคำถามข้อใดเด็กเก่งมักตอบผิดเป็นจำนวนมากกว่าเด็กอ่อน อำนาจจำแนกของข้อนั้นก็จะเป็นลบ. วิธีสถิติที่ใช้สำหรับหาอำนาจจำแนกของข้อคำถามก็มีหลายแบบ แต่แบบที่นิยมและยอมรับกันทั่วไปในวงการวิจัยคือ วิธีหาสหสัมพันธ์แบบ Biserial correlation หรือ Point - biserial correlation ($r_{p.bis}$) ดังสูตรที่แสดงไว้ในภาคผนวก ก.

สำหรับการหาอำนาจจำแนกของแบบทดสอบทั้งฉบับของรายงานนี้ ใช้วิธีสถิติแบบ Serial correlation (r_{ser}) ดังสูตรที่แสดงไว้ในภาคผนวก ก. หลักการของวิธีวิเคราะห์ที่ปรับปรุงมาจากวิธี Point biserial แบบสามัญ ซึ่งเป็นการหาสหสัมพันธ์ระหว่างตัวแปรที่มีการแจกแจงต่อเนื่องกัน (Continuous variable) กับตัวแปรอีกฝ่ายหนึ่งที่แจกได้เป็น 2 กลุ่มหรือ 2 ประเภท. ส่วนการหาอำนาจจำแนกของรายงานนี้เป็นการหาสหสัมพันธ์ระหว่างคะแนนที่เด็กสอบได้ในแต่ละฉบับ ซึ่งถือว่าเป็นตัวแปรต่อเนื่อง กับตัวแปรอีกฝ่ายหนึ่งที่มีการแจกแจงเป็นสามชั้นเรียน (trichotomy) และกับตัวแปรที่แจกออกเป็น 9 ประเภทอายุ. ฉะนั้น ในการวิเคราะห์จึงต้องใช้เทคนิคแบบ Serial correlation มาทำการวิจัย. ตามวิเคราะห์ปรากฏว่า r_{ser} ที่มีความเชื่อมั่นสูงอย่างมีนัยสำคัญทางสถิติแล้ว ก็หมายความว่า แบบทดสอบนี้มีอำนาจจำแนกนักเรียนออกเป็นประเภท ๆ ตามนั้นได้จริง เป็นแบบทดสอบที่มีคุณภาพเหมาะที่จะใช้เป็นเครื่องมือสำหรับการวัดผลได้ ซึ่งก็จะให้ผลต่อเนื่องกับการวิจัยว่า แบบทดสอบประเภทใดจะเหมาะสมกับนักเรียนเหล่านั้นมากที่สุด.

ในการวิเคราะห์นี้ จะได้นำเสนอแต่เพียงผลรวบยอดของการคำนวณเท่านั้น รายละเอียดของตัวเลขข้อมูลระหว่างการคำนวณเพื่อจะได้ผลลัพธ์รวบยอดนั้น ได้แสดงไว้แล้วในภาคผนวก ข.

ข. 1 อำนาจจำแนกของแบบทดสอบโดยชั้นเรียน

การวิเคราะห์หนึ่งที่จะคนกว่า แบบทดสอบแต่ละฉบับมีอำนาจจำแนกแตกต่างกันเป็น 3 ระดับชั้นเรียน ดังที่ได้อธิบายไว้หรือไม่ ถือว่าเป็นนักเรียนชั้นต่ำที่จะตอบได้คะแนนน้อย และชั้นสูงก็จะตอบได้คะแนนมากขึ้นตามลำดับ ดังผลการวิเคราะห์ที่แสดงไว้ในตารางข้างล่าง.

ตาราง 26. ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามระดับชั้นเรียน

แบบทดสอบ	r_{ser}	r_{ser}	t
อุปมาอุปไมย-ตัวหนังสือ	.750	$\pm .0121$	41.0688 **
อุปมาอุปไมย-ภาพ	.756	$\pm .0118$	41.8287 **
ไม่เขาพวก-ตัวหนังสือ	.765	$\pm .0114$	56.2425 **
ไม่เขาพวก-ภาพ	.660	$\pm .0156$	48.2384 **

ตารางนี้แสดงว่า แบบทดสอบทั้ง 4 ฉบับ ต่างมีอำนาจจำแนกสูงเป็นที่พอใจเกินกว่าระดับ .75 ยกเว้นแบบทดสอบไม่เขาพวกประเภทภาพเพียงฉบับเดียวที่มีอำนาจจำแนก .66 และเป็นที่น่ายินดีว่า ค่าอำนาจการจำแนกของแบบทดสอบเหล่านี้ มีความคลาดเคลื่อนน้อยมาก ในระดับ $\pm .01$ เป็นส่วนใหญ่ นั่นคืออำนาจจำแนกของแบบทดสอบอุปมาอุปไมยที่ใช้ตัวหนังสือ จะมีค่าอำนาจจำแนกจริง ในระดับความเชื่อมั่น 99 % ระหว่าง .719 ถึง .781 ส่วนค่าของ t ในช่องสุดท้ายของตารางเป็นตัวเลขที่ทดสอบว่า แต่ละค่า r_{ser} ที่คำนวณได้จะเชื่อมั่นได้มากน้อยเพียงใด ซึ่งก็เป็นที่ยืนยันที่ผลการทดสอบยืนยันว่า ค่าอำนาจจำแนกเหล่านี้มีค่าสูงมากจนไว้วางใจได้ ดังนั้นเราอาจสรุปได้ว่า นักเรียนแต่ละชั้นที่ทำการทดลองนี้มีระดับความสามารถในการทำแบบทดสอบเหล่านี้แตกต่างกันจริง นั่นคือ แบบทดสอบทั้ง 4 ฉบับต่างมีอำนาจจำแนกสามารถแยกแยะว่า ผู้สอบคนใดมีระดับความสามารถมากน้อยแตกต่างกันจริง.

ข. 2 อำนาจจำแนกของแบบทดสอบโดยอายุ

เนื่องจากการทดลองนี้ ได้กระทำกับเด็กนักเรียน 9 ระดับอายุจาก 7 ขวบ จนถึง 15 ขวบ ฉะนั้น จึงเห็นสมควรที่จะตรวจสอบว่า แบบทดสอบเหล่านี้จะมีอำนาจจำแนกสามารถแยกเด็กออกเป็นแต่ละระดับอายุตามนั้นได้หรือไม่ อีกประการหนึ่งด้วย ดังที่แสดงผลการวิเคราะห์ไว้ในตารางข้างล่างนี้

ตาราง 27. ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามระดับอายุ

แบบทดสอบ	r_{ser}	$\sigma_{r_{ser}}$	t
อุปมาอุปไมย-ตัวหนังสือ	.478	$\pm .0213$	19.7331**
อุปมาอุปไมย-ภาพ	.482	$\pm .0212$	19.9205**
ไม่เขาพวก-ตัวหนังสือ	.493	$\pm .0209$	20.5405**
ไม่เขาพวก-ภาพ	.468	$\pm .0215$	19.1841**

ผลการวิเคราะห์ปรากฏว่า แบบทดสอบทั้ง 4 ฉบับต่างมีอำนาจจำแนกสูงพอที่จะแยกเด็กออกจากตามระดับอายุเป็นที่เชื่อมั่นได้เกิน 99 % และค่าของ r_{ser} ที่หาได้นั้นมีความคลาดเคลื่อนน้อยมากในระดับ .021 เท่านั้น. ตัวเลขเหล่านี้ แสดงให้เห็นคุณภาพของแบบทดสอบทั้ง 4 ชนิดว่าอยู่ในเกณฑ์ดี สามารถชั่งนอกว่าผู้สอบที่มีอายุน้อยก็มักจะได้คะแนนน้อยกว่าผู้ที่มีอายุมาก.

เป็นที่น่าสังเกตว่า คุณภาพในการจำแนกโดยอายุนี้มีค่า r_{ser} น้อยกว่าเมื่อจำแนกโดยชั้นอย่างเด่นชัด. ทั้งนี้เนื่องจากเป็นการหาค่า r_{ser} ที่ตัดออกเป็นกลุ่มย่อย ๆ ที่มีความสามารถใกล้เคียงกันมาก. แต่อย่างไรก็ดี เราก็อาจสรุปได้ว่า คุณภาพของแบบทดสอบทั้ง 4 ฉบับนี้จะมีคุณภาพค่อยลงเมื่อนำไปหาอำนาจจำแนกที่แจ่มแจ้งเป็นรายอายุ.

ข. 3 อำนาจจำแนกของแบบทดสอบโดยเพศ

โดยทำนองเดียวกันกับการวิเคราะห์ใน 2 ตารางก่อน ได้ตัดนักเรียนออกเป็น 2 กลุ่มตามเพศ แล้วมาคำนวณหาอำนาจจำแนกและอื่น ๆ ดังแสดงไว้ในตารางที่ 28.

ตาราง 28. ค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับตามเพศ

แบบทดสอบ	$r_{p.bis}$	$r_{p.bis}$	t
อุปมาอุปไมย-ตัวหนังสือ	.092	$\pm .0274$	2.6535*
อุปมาอุปไมย-ภาพ	.107	$\pm .0256$	10.1414**
ไม่เข้าใจ-ตัวหนังสือ	.123	$\pm .0273$	3.5875**
ไม่เข้าใจ-ภาพ	.132	$\pm .0273$	3.8405**

เนื่องจากการหาอำนาจจำแนกครั้งนี้เป็นการกระทำกับตัวแปรที่แยกได้เป็นเพียง 2 ลักษณะ (dichotomy) เราจึงสามารถหาสหสัมพันธ์โดยวิธี Point biserial แบบธรรมดา ซึ่งก็ปรากฏได้ว่าค่า $r_{p.bis}$ ยิ่งต่ำลงไปมาก แต่ก็ยังอยู่ในระดับที่เชื่อมั่นได้โดยทุกฉบับ โดยเฉพาะฉบับแรกซึ่งเป็นอุปมาอุปไมยประเภทตัวหนังสือจะเชื่อมั่นได้เพียงระดับ 5 % เท่านั้น. ฉะนั้นเราจึงสรุปได้ว่า แบบทดสอบชุดนี้มีอำนาจจำแนกต่ำในการที่จะแยกว่า ผู้สอบคนใดเป็นเพศชายหรือหญิง.

เมื่อนำผลวิเคราะห์จากตารางที่ 26 - 28 มาเปรียบเทียบกัน ก็จะเห็นได้ว่าแบบทดสอบทั้ง 4 ฉบับนี้ต่างจะมีค่าอำนาจการจำแนกแต่ละลักษณะแตกต่างกัน. อำนาจจำแนกแตกต่างกันตามลักษณะของการแยกกลุ่ม. ถ้าเราแยกกลุ่มโดยถือชั้นเรียนเป็นเกณฑ์ อำนาจจำแนกจะมีค่าสูงที่สุดถึงระดับ .75 ถ้าแยกกลุ่มโดยถืออายุเป็นเกณฑ์จะลดเหลือประมาณ .48 และถ้าแยกโดยถือเพศจะมีค่าอำนาจจำแนกเพียง .12 สาเหตุที่เป็นเช่นนี้ อาจแปลได้เป็น 2 นัยคือ ในประการแรกนั้น ระดับความสามารถของเด็กในแต่ละชั้นจาก ป. 2 ไป ป.4 และ ป.7 นั้นแตกต่างกันอย่างเด่นชัด ส่วนความสามารถของเด็กที่มีอายุเพิ่มขึ้นแต่ละปี ๆ นั้น ปริมาณน้อยกว่าประเภทแรก จึงมีค่าอำนาจจำแนก

คำว่า ส่วนการแบ่งกลุ่มโดยเพศนั้นเกือบไม่มีความหมายเลย เพราะทั้งนักเรียนชายและหญิงต่างก็มีความสามารถในเรื่องนี้ปาน ๆ กัน จึงให้ค่าอำนาจจำแนกน้อยที่สุด.

หรือในประการที่สองนั้น เราอาจมองกลับไปค้นคุณภาพของแบบทดสอบบ้างก็ได้ ว่าแบบทดสอบนี้มีความไว ในการจำแนกเด็กเก่ง - อ่อนของแต่ละชั้นเรียนได้ดีที่สุด. ส่วนความสามารถที่จะจำแนกเด็กเก่งอ่อนภายในแต่ละอายุชั้นจะมีความไวลดน้อยลง และถ้าจะจำแนกเป็นเด็กเก่ง - อ่อนในแต่ละเพศแล้ว จะมีความไวน้อยที่สุด.

จากผลวิเคราะห์ดังกล่าวมานี้ จึงจะใคร่ทราบต่อไปอีกชั้นหนึ่งว่า ในแบบทดสอบทั้ง 4 ฉบับนี้ ฉบับใดจะส่งผลกระทบต่อคะแนนรวมของเด็กแต่ละประเภทมากน้อยเพียงใด ดังจะได้กล่าวในตอนต่อไป

ค. สมการพยากรณ์ตามระดับชั้นเรียนและอายุ

ในการวิเคราะห์ขั้นสุดท้ายนี้ จะได้นำคะแนนผลการสอบของนักเรียนทั้งหมดมาแยกเป็นรายชั้นและรายอายุ ดังที่เคยกระทำมาแล้ว. แต่คราวนี้ จะได้นำคะแนนเหล่านั้นมาวิเคราะห์หา Multiple serial correlation โดยวิธี Discriminant function ผลที่ได้จากการคำนวณนี้จะออกมาเป็นรูปของสมการ ซึ่งสามารถนำไปใช้ในเชิงพยากรณ์ว่า ผู้สอบคนนั้นมีความสามารถเทียบเท่ากับเด็กนักเรียนระดับชั้นใด หรือเทียบเท่ากับนักเรียนระดับอายุใดก็ได้ และอีกประการหนึ่งนั้น สมการนี้จะสามารถบอกให้เราทราบอีกด้วย ว่าแบบทดสอบฉบับใดมีอิทธิพลมากน้อย ต่อผลการสอบของนักเรียนแต่ละประเภท ๆ นั้นด้วย. การที่เราทราบน้ำหนักหรือสัดส่วนของแต่ละแบบทดสอบที่มีต่อการสอบนี้ ก็เท่ากับเป็นการบอกให้เราทราบโดยปริยายว่า เด็กนักเรียนระดับชั้นใดหรืออายุใด ต่างจะมีสมรรถภาพชนิดใดมากน้อยเป็นสัดส่วนเท่าใดนั่นเอง. ดังจะได้เสนอผลวิเคราะห์ต่อไปโดยลำดับดังนี้

ค. 1 สมการสำหรับพยากรณ์ระดับชั้นเรียน

ขบวนการหาสมการ Multiple serial regression นี้จะหาจากสูตรสหสัมพันธ์พหุคูณของ Point serial correlation (R_p) ได้เสนอไว้แล้วในภาคผนวก ก.

$$R_p = \sqrt{\frac{\Delta}{N \sum \left[\frac{(y_1 - y_h)^2}{p} \right]}}$$

ในการคำนวณจะเริ่มด้วยการหา Ordinate 2 เส้น (y) ที่แบ่งนักเรียนทั้งหมดออกเป็น 3 ส่วน (p) ตามระดับชั้นประถมศึกษาปีที่ 2, ปีที่ 4, และ ปีที่ 7 คำนวณค่าต่าง ๆ ที่แสดงไว้ในตารางที่ 29 จากนั้นจะได้อัตรา Δ ตัวแปรซึ่งหมายถึงผลรวมของ regression ของแบบทดสอบแต่ละฉบับกำลังสองจากการแกสมการหลายชั้นเพื่อหาสัมประสิทธิ์ของ triserial correlation กำลังสมการข้างล่าง



ตาราง 29. การหา R_p และสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับตามระดับชั้นเรียน

ชั้น เรียน	N	P	y	yl-yh	$\frac{yl-yh}{p}$	$\frac{(yl-yh)^2}{p}$	ผลรวม			
							$\sum x_1$	$\sum x_2$	$\sum x_3$	$\sum x_4$
ป. 2	424	0.3222	0.35863	-0.35863	1.11306	0.3991817900	5037	614	5059	7718
			0.35863	0.00032	.00090	0.0000000028	7823	9250	8109	10825
ป. 4	467	0.3549	0.35895	0.38595	1.19526	0.3990247800	10577	11306	10563	11587
ป. 7	425	0.3229								
Total	1316	1.0000				0.7982046000	23437	26670	23731	30130

สมการถดถอยสำหรับหาสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับ

$$\begin{aligned}
 7042.8225 &= 71663.1269a_1 + 59582.7554a_2 + 75252.2014a_3 + 14758.6794a_4 \\
 6716.6857 &= 59582.7554a_1 + 61942.9149a_2 + 42537.5639a_3 + 42449.7022a_4 \\
 7001.8589 &= 75252.2014a_1 + 42537.5639a_2 + 67759.5950a_3 + 41402.5092a_4 \\
 5268.6230 &= 14758.6794a_1 + 42449.7022a_2 + 41402.5092a_3 + 46205.5289a_4
 \end{aligned}$$

นั่นคือ $a_1 = 0.01503$, $a_2 = 0.04427$, $a_3 = 0.03748$, $a_4 = 0.03497$

และ $R_p = .921$; $F = 1388.59$

สหสัมพันธ์พหุคูณของการพยากรณ์โดยชั้น ปรากฏว่ามีความสูงถึง .921 โดยมีระดับความเชื่อมั่นเกินกว่า 99 % และอาจเขียนเป็นรูปสมการสำหรับพยากรณ์แบบคะแนนมาตรฐาน (z score) และคะแนนดิบ ดังนี้

$$\tilde{v} = .01503 x_1 + .04427 x_2 + .03748 x_3 + .03497 x_4$$

ในเมื่อ $\tilde{v}$ แทน ระดับชั้นในหน่วยของคะแนนมาตรฐาน

x_1 แทน คะแนนมาตรฐานของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ

x_2 แทน คะแนนมาตรฐานของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

x_3 แทน คะแนนมาตรฐานของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ

x_4 แทน คะแนนมาตรฐานของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

เนื่องจากสูตรคะแนนมาตรฐานต้องการนำไปใช้ในเชิงปฏิบัติ จึงได้แปลงสูตร เบื้องต้นให้เป็นหน่วยของคะแนนดิบ ดังสมการนี้

$$\tilde{v} = .01503 x_1 + .04427 x_2 + .03748 x_3 + .03497 x_4 + 1.79910$$

ในเมื่อ $\tilde{v}$ แทน ระดับชั้นในหน่วยของคะแนนมาตรฐาน

x_1 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ

x_2 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

x_3 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ

x_4 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

สมการพยากรณ์ชั้นเรียนชั้นต้นมีความหมายว่า ถ้านักเรียนคนใดสอบได้คะแนนจากแต่ละแบบทดสอบทั้ง 4 ฉบับเป็นเท่าใดแล้ว ถ้าเรานำคะแนนนั้นไปคูณกับสัมประสิทธิ์ของแต่ละแบบทดสอบเป็นคู่ ๆ กันไป ก็จะได้ตัวเลขสำหรับพยากรณ์ว่า เด็กนั้นมีความรู้ เทียบเสมือนกับ เด็กชั้นเรียนใด

อนึ่ง ตัวเลขสัมประสิทธิ์ของแต่ละแบบทดสอบนั้น ก็หมายถึงน้ำหนักหรืออิทธิพลที่แบบทดสอบนั้นส่งผลกระทบต่อชั้นเรียน หรือจะพูดกลับเพื่อให้เห็นง่ายขึ้นก็ได้ว่า เด็กนักเรียนแต่ละชั้นจะมีสมรรถภาพอย่างละเท่าใด ก็ได้. และเพื่อจะเทียบให้เห็นว่า แบบทดสอบแต่ละฉบับต่างจะส่งผลกระทบต่อชั้นเรียนเป็นสัดส่วนเท่าไร จึงได้คำนวณให้เป็นจำนวนร้อยละได้ ดังนี้

แบบทดสอบอุปมาอุปไมย-ตัวหนังสือ	ส่งผล .01503 หรือเท่ากับ	11.27 %
แบบทดสอบอุปมาอุปไมย-รูปภาพ	ส่งผล .04427 หรือเท่ากับ	33.20 %
แบบทดสอบไมเข้าพวก-ตัวหนังสือ	ส่งผล .03748 หรือเท่ากับ	28.11 %
แบบทดสอบไมเข้าพวก-รูปภาพ	ส่งผล .03497 หรือเท่ากับ	26.23 %

ตัวเลขนี้แสดงว่า แบบทดสอบอุปมาอุปไมยประเภทรูปภาพ ส่งผลต่อการพยากรณ์เรียน หรือมีอิทธิพลที่จะคลี่ให้นักเรียนอยู่ในชั้นเรียนใด เป็นจำนวนสูงมากที่สุดถึงร้อยละ 33. ที่มีอิทธิพล รองลงมาคือแบบทดสอบไมเข้าพวกประเภทตัวหนังสือ 28 %, ไมเข้าพวกรูปภาพ 26 % และ อุปมาอุปไมยตัวหนังสือประมาณ 11 % ฉะนั้น เด็กคนใดที่มีความสามารถสูงในอุปมาอุปไมยรูปภาพ จะมีโอกาสเป็นเด็กที่อยู่ในชั้นเรียนสูง มากกว่าผู้ที่มีความสามารถในแบบทดสอบชนิดอื่น ๆ และเป็นที่ น่าสังเกตว่า เฉพาะแบบทดสอบประเภทรูปภาพทั้ง 2 ชนิดนั้น มีอิทธิพลรวมกันถึง 60 % มากกว่า อีก 2 ชนิดที่ใช้ตัวหนังสือ. นี่แสดงว่า แบบทดสอบประเภทรูปภาพมีความสามารถในการพยากรณ์ หรือมีอิทธิพลต่อระดับชั้นเรียนมากกว่าผู้ทบทองทางภาษา.

ค. 2 สมการสำหรับพยากรณ์ตามระดับอายุ

โดยวิธีการคำนวณแบบเดียวกันกับชอกอน ได้ทำการแจกแจงของนักเรียนทั้งหมดมาจัดแบ่ง เป็น 9 กลุ่ม ตามสัดส่วนของนักเรียนแต่ละอายุ. จากนั้นได้คำนวณค่าสถิติต่าง ๆ เพื่อนำ ไปหา R_p และสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับจาก a_1 ถึง a_4 ดังตารางข้างล่างนี้

ตาราง 30. การหา R_p และสัมประสิทธิ์ของแบบทดสอบแต่ละฉบับตามระดับอายุ

อายุ	N	p	y	yl-yh	$\frac{yl-yh}{p}$	$\frac{(yl-yh)^2}{p}$	ผลบวก			
							$\sum x_1$	$\sum x_2$	$\sum x_3$	$\sum x_4$
7	74	0.05623	0.11320	-0.11320	-2.01316	0.22834	868	1072	846	1235
8	154	0.11702	0.25613	-0.14293	-1.22141	0.17458	1891	2268	1879	2827
9	157	0.11930	0.34371	-0.08758	-0.73411	0.06429	2175	2639	2294	3187
10	208	0.15805	0.39588	-0.05217	-0.33008	0.01720	3297	3916	3427	4664
13	211	0.16033	0.38341	0.01247	0.07777	0.00093	5019	5350	4974	5564
11	189	0.14361	0.31466	0.06875	0.47872	0.03293	3093	3604	3115	4271
12	157	0.11930	0.20701	0.10765	0.90234	0.09715	3129	3525	3205	3918
14	108	0.08206	0.09321	0.11380	1.38679	0.15781	2628	2796	2590	2885
15	58	0.04407	0.09321	0.09321	2.11504	0.19718	1339	1500	1401	1579
Total	1316	1.00000				0.97041				

จากตารางนี้ สัมประสิทธิ์การหาค่าตามข้อได้แก่

$$\begin{aligned}
 4430.3491 &= 71663.1269a_1 + 59582.7554a_2 + 75252.2014a_3 + 14758.6794a_4 \\
 4213.9747 &= 59582.7554a_1 + 61942.9149a_2 + 42537.5639a_3 + 42449.7022a_4 \\
 4511.6024 &= 75252.2014a_1 + 42537.5639a_2 + 67759.5950a_3 + 41402.5092a_4 \\
 3534.9505 &= 14758.6794a_1 + 42499.7022a_2 + 41402.5092a_3 + 46205.5289a_4
 \end{aligned}$$

$$\begin{aligned}
 \text{นั่นคือ } a_1 &= .00577, a_2 = .02494, a_3 = .02850, a_4 = .02621 \\
 \text{และ } R_p &= .533; F = 124.66100
 \end{aligned}$$

จากตารางที่ 30 ปรากฏว่า สหสัมพันธ์พหุคูณของการพยากรณ์ตามระดับอายุมีค่า .533 ซึ่งน้อยกว่าสหสัมพันธ์พหุคูณของการพยากรณ์โดยชั้นเรียน แต่อย่างไรก็ตาม สหสัมพันธ์พหุคูณก็ยังมีระดับความเชื่อมั่น 99 % และสมการพยากรณ์ก็เช่นเดียวกันกับข้อ ค. 1 คือเขียนได้ 2 สมการคือ สมการรูปคะแนนมาตรฐาน เขียนได้ดังนี้

$$\bar{Y} = .00577 x_1 + .02494 x_2 + .02850 x_3 + .02121 x_4$$

เมื่อ $\bar{Y}$ แทน ระดับอายุในหน่วยของคะแนนมาตรฐาน

x_1 แทน คะแนนมาตรฐานของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ

x_2 แทน คะแนนมาตรฐานของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

x_3 แทน คะแนนมาตรฐานของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ

x_4 แทน คะแนนมาตรฐานของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

อีกสมการหนึ่งในรูปสมการคะแนนดิบ คือ

$$\bar{Y} = .00577 X_1 + .02494 X_2 + .02850 X_3 + .02621 X_4 + 2.50777$$

ในเมื่อ $\bar{Y}$ แทน ระดับอายุในหน่วยของคะแนนมาตรฐาน

X_1 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ

X_2 แทน คะแนนดิบของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

X_3 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ

X_4 แทน คะแนนดิบของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

จากสมการพยากรณ์ที่ 2 สมการนี้ หมายความว่า ถ้าเด็กคนใดสอบได้คะแนนเท่าใดก็ไปแทนค่าในสมการ ก็จะทราบว่า นักเรียนคนนั้นมีความรู้ เทียบเท่ากับเด็กระดับอายุใดใด และน้ำหนักของแบบทดสอบแต่ละฉบับจะส่งผลในแต่ละระดับอายุเป็นสัดส่วนกัน ดังนี้

แบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือส่งผล 6.75 %

แบบทดสอบอุปมาอุปไมยประเภทรูปภาพส่งผล 29.34 %

แบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ 33.53 %

แบบทดสอบไม่เข้าพวกประเภทรูปภาพ 30.84 %

จากจำนวนเปอร์เซ็นต์ข้างบนจะเห็นว่า แบบทดสอบที่ส่งผลมากที่สุด คือแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ รองลงมาคือแบบทดสอบไม่เข้าพวกประเภทรูปภาพ และแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือส่งผลน้อยมาก แต่ก็ยังสามารถเป็นตัวพยากรณ์ได้ในสมการพยากรณ์ระดับอายุนี้.

เมื่อนำสัมประสิทธิ์ของสมการพยากรณ์ตามระดับชั้น และตามระดับอายุมาเปรียบเทียบกัน ก็จะเห็นว่า แบบทดสอบอุปมาอุปไมยประเภทรูปภาพส่งผลการพยากรณ์ชั้นเรียนสูงมากถึง 33 % เป็นอันดับหนึ่ง และแบบทดสอบไม่เข้าพวกตัวหนังสือส่งผล 28 % เป็นอันดับที่สอง. แต่ในการพยากรณ์ตามอายุนั้น จะปรากฏว่า แบบทดสอบไม่เข้าพวกประเภทตัวหนังสือจะส่งผลมากเป็นอันดับที่หนึ่งถึง 34 % แบบทดสอบไม่เข้าพวกประเภทรูปภาพจะส่งผลเป็นรองลงมา นั่นคือ แบบทดสอบไม่เข้าพวกทั้งสองประเภทมีคุณภาพสูงในคำพยากรณ์อายุผู้สอบ. สำหรับแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือนั้นส่งผลการพยากรณ์ทั้งชั้นเรียนและอายุเป็นอันดับสุดท้าย

รายงานนี้มิได้สร้างสมการสำหรับพยากรณ์โดยเพศชนเลย เพราะจากผลวิเคราะห์ในตารางที่ 28 หน้า 66 นั้น ปรากฏว่า สัมประสิทธิ์ของอำนาจจำแนกตามเพศมีค่าต่ำมาก อันแสดงว่า เพศไม่มีอิทธิพลต่อคะแนนการสอบอย่างเด่นชัด ทั้งเพศชายและเพศหญิงต่างจะได้คะแนนปาน ๆ กัน จึงไม่มีเหตุผลสมควรที่จะสร้างสมการพยากรณ์

อนึ่ง ในการนำสมการพยากรณ์ในรูปคะแนนมาตรฐานหรือคะแนนดิบไปใช้สำหรับทำนายระดับชั้นเรียนหรืออายุนั้น หลังจากทราบค่า $\tilde{z}$ หรือ $\tilde{v}$ ซึ่งบอกเป็นหน่วยของคะแนนมาตรฐาน (z) แล้ว - จะต้องนำ $\tilde{z}$ นั้นไปอ่านจากตารางพื้นที่โค้งปกติให้เป็นสัดส่วนเสียก่อนว่า ค่า $\tilde{z}$ หรือ $\tilde{v}$ นั้นตกอยู่ในขอบเขตของสัดส่วนชั้นใด หรืออายุใด ดังแสดงไว้ในตารางที่ 29 และ 30 ก็พยากรณ์ชั้นเรียนและอายุของผู้สอบไปตามนั้น ๆ.

สรุปผลการวิเคราะห์ และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะเปรียบเทียบแบบทดสอบที่เป็นตัวหนังสือ และแบบทดสอบที่เป็นรูปภาพว่า แบบทดสอบชนิดใดและประเภทใดที่เหมาะสมที่สุดในระดับชั้นประถมศึกษา และเหมาะสมกับระดับชั้นเรียนใด อายุใด และเพศใดมากที่สุด อีกประการหนึ่งผู้วิจัยต้องการจะสร้างแบบทดสอบอุปมาอุปไมยกับแบบทดสอบไม่เขาพวกทั้งประเภทตัวหนังสือและรูปภาพ เพื่อใช้ในการพยากรณ์ระดับชั้นเรียนและระดับอายุของนักเรียน เมื่อเด็กเหล่านั้นได้ถูกทดสอบไปแล้ว.

กลุ่มตัวอย่างที่ใช้ในการศึกษา

การศึกษานี้ศึกษาจากนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4 และปีที่ 7 จากโรงเรียน
ทั้ง 12 แห่งคือ

1. โรงเรียนประจักษ์ศิลปาคม
2. โรงเรียนประจักษ์ศิลปาคม
3. โรงเรียนวัดธาตุทอง
4. โรงเรียนวัดลาดบัวขาว
5. โรงเรียนสวนกุหลาบ
6. โรงเรียนวิบูลย์
7. โรงเรียนวัดชนะสงคราม
8. โรงเรียนทุ่งมหาเมฆ
9. โรงเรียนลาซาล
10. โรงเรียนอัสสัมชัญ
11. โรงเรียนพรวิทยา
12. โรงเรียนครุฑนิเวศน์

โรงเรียนทั้ง 12 แห่งนี้ มีจำนวนนักเรียนที่ทำการศึกษา 1316 คน เป็นชั้นประถมศึกษาปีที่ 2 จำนวน 424 คน, ชั้นประถมศึกษาปีที่ 4 จำนวน 467 คน และชั้นประถมศึกษาปีที่ 7 จำนวน 425 คน.

เครื่องมือในการเก็บข้อมูล

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบความถนัดทางการเรียน 4 ฉบับ ซึ่งเป็นแบบทดสอบปรนัยชนิดเลือกตอบ มี 5 ตัวเลือก และแบ่งเป็น 4 ฉบับ ดังนี้

1. แบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ เป็นแบบทดสอบปรนัยชนิดเลือกตอบมี 40 ข้อ กำหนดเวลาให้ทำ 20 นาที มีความยากมาตรฐานเฉลี่ยเป็นรายข้อ (Δ) 13.0525 และมีความเชื่อมั่น .855

2. แบบทดสอบอุปมาอุปไมยประเภทรูปภาพ เป็นแบบทดสอบปรนัย ชนิดเลือกตอบมี 40 ข้อ กำหนดเวลาให้ทำ 20 นาที มีความยากมาตรฐานเฉลี่ย (Δ) 13.0525 และมีความเชื่อมั่น .855

3. แบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ เป็นแบบทดสอบปรนัยชนิดเลือกตอบ กำหนดเวลาให้ทำ 20 นาที มีความยากมาตรฐานเฉลี่ยเป็นรายข้อ (Δ) 13.2350 และมีความเชื่อมั่น .829

4. แบบทดสอบไม่เข้าพวกประเภทรูปภาพ เป็นแบบทดสอบปรนัยชนิดเลือกตอบ กำหนดเวลาให้ทำ 20 นาที มีความยากมาตรฐานเฉลี่ยเป็นรายข้อ (Δ) 13.2350 และมีความเชื่อมั่น .740

แบบทดสอบทั้ง 4 ฉบับนี้ ยังใช้เป็นตัวพยากรณ์ ในการทำนายระดับชั้นเรียนและระดับอายุ ของนักเรียนเช่นกัน.

9. การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลจากการศึกษาครั้งนี้ ได้อาศัยเครื่องคำนวณฟรีเดน (Friden) ทำสิ่งต่อไปนี้

1. หาผลรวมทั้งหมดของคะแนน
2. หาผลรวมของคะแนนแต่ละตัวยกกำลังสอง
3. หาคะแนนเฉลี่ย
4. หาความเบี่ยงเบนมาตรฐาน
5. หาคะแนนมาตรฐาน
6. หาสหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบ และสหสัมพันธ์ภายในของอำนาจจำแนกของแบบทดสอบ
7. ทดสอบนัยสำคัญของความแตกต่างของคะแนนเฉลี่ย โดยวิธี t - test และ Analysis of variance
8. หาสมการพยากรณ์ระดับชั้นเรียน และระดับอายุ

10. ผลจากการวิเคราะห์ข้อมูล

ผลจากการวิจัยทำให้พบสิ่งต่อไปนี้คือ

1. จากคะแนนความแตกต่างของนักเรียนในระดับประถมศึกษา ปรากฏว่า แบบทดสอบประเภทรูปภาพได้คะแนนเฉลี่ยเป็นครั้งแรกและมากกว่าเล็กน้อยของคะแนนเต็ม และคะแนนเฉลี่ยจากแบบทดสอบประเภทรูปภาพ จะมากกว่าคะแนนเฉลี่ยของแบบทดสอบประเภทตัวหนังสือเสมอ และแตกต่างกันอย่างมีนัยสำคัญเชิงสถิติด้วย.

ส่วนในด้านการเชื่อมั่นของแบบทดสอบนี้ แบบทดสอบทั้งประเภทตัวหนังสือและภาพต่างมีความเชื่อมั่นไม่ผิดแผกจากกันเลย และต่างก็มีความคลาดเคลื่อนมาตรฐานของคะแนนน้อย

2. คะแนนเฉลี่ยที่แยกตามระดับชั้นเรียนของแบบทดสอบทั้ง 4 ฉบับ จะสูงขึ้นเรื่อย ๆ ตามระดับชั้นและคะแนนของแบบทดสอบประเภทรูปภาพจะมากกว่าคะแนนเฉลี่ยของแบบทดสอบประเภทตัวหนังสือในทุก ๆ ระดับชั้นเรียนอย่างมีนัยใจได้.

ในด้านคุณภาพของแบบทดสอบนั้น มีค่าความเชื่อมั่นสูงในแต่ละระดับชั้นเรียน และมีค่าอำนาจจำแนกสูงกว่า .7 ขึ้นไปเว้นแต่แบบทดสอบไม่เข้าพวกประเภทรูปภาพ ก็มีค่าอำนาจจำแนก .66 แต่ก็ยังนับว่าเป็นแบบทดสอบที่เชื่อมั่นได้ในด้านการแยกเด็กเก่ง - อ่อน เป็นต้น.

จากผลที่กล่าวข้างต้น อาจสรุปได้ว่า ^①ชั้นประถมศึกษาปีที่ 2, เหมาะที่จะใช้แบบทดสอบประเภทรูปภาพมากที่สุด โดยเฉพาะแบบทดสอบไม่เข้าพวกประเภทรูปภาพ ก็มีคะแนนเฉลี่ย 18.2028 ซึ่งเป็นคะแนนเฉลี่ยที่ใกล้เคียงกับคะแนนเต็ม และถ้าเราสังเกตคะแนนเฉลี่ยในรูปของคะแนน T^② ประสิทธิภาพของแบบทดสอบทั้ง 4 ฉบับแล้ว จะเห็นว่า ^③แบบทดสอบทั้ง 4 ฉบับมีความยากง่ายเหมาะสมกับนักเรียนชั้นประถมศึกษาปีที่ 4 มากที่สุด เพราะคะแนนเฉลี่ยของทุกฉบับมีค่าเท่ากับหรือใกล้เคียงกับ T 50 ทั้งสิ้น ^④ส่วนแบบทดสอบที่เหมาะสมกับชั้นประถมศึกษาปีที่ 7 คือแบบทดสอบประเภทตัวหนังสือทั้ง 2 ชนิด เพราะประเภทรูปภาพจะง่ายเกินไปสำหรับในระดับนี้.

3. จากคะแนนความแตกต่างระหว่างอายุ ปรากฏว่า คะแนนเฉลี่ยของแบบทดสอบแต่ละฉบับจะเพิ่มขึ้นตามระดับอายุ เช่นเดียวกับกับคะแนนเฉลี่ยของแต่ละชั้นเรียน และแบบทดสอบที่เหมาะสมที่สุดในเด็กอายุ 12 ปีคือ แบบทดสอบรูปภาพไม่เข้าพวกประเภทตัวหนังสือและรูปภาพและแบบทดสอบไม่เข้าพวกตัวหนังสือ ส่วนแบบทดสอบไม่เข้าพวกประเภทรูปภาพเหมาะสมกับเด็กอายุ 9 ปีเพราะอายุเหล่านี้นักเรียนมีคะแนนเฉลี่ย และค่าความเชื่อมั่น ของแบบทดสอบแต่ละฉบับสูงที่สุด.

ในแง่ความเชื่อมั่นของแบบทดสอบแล้ว อายุต่ำจะมีค่าความเชื่อมั่นสูงในแบบทดสอบประเภทรูปภาพ ส่วนอายุมากก็มีค่าความเชื่อมั่นของแบบทดสอบประเภทตัวหนังสือสูง นั่นคือแบบทดสอบประเภทรูปภาพเหมาะสมกับระดับเด็กอายุต่ำ ๆ และแบบทดสอบประเภทตัวหนังสือเหมาะสมกับเด็กอายุมาก ๆ.

ส่วนค่าอำนาจจำแนกของแบบทดสอบทั้ง 4 ฉบับที่แยกตามระดับอายุ จะมีค่าต่ำกว่าอำนาจจำแนกที่แยกตามระดับชั้น ทั้งนี้เพราะการจำแนกตามอายุจะมีนักเรียนที่มีความสามารถใกล้เคียงกันในแต่ละอายุ ๆ จึงทำให้ค่าอำนาจจำแนกในการแยกเด็กเก่ง - อ่อนต่ำลง.

4. คะแนนเฉลี่ยที่ได้จากแบบทดสอบที่แยกโดยเพศ จะใกล้เคียงกัน ไม่ค่อยมีคะแนนเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญนัก ส่วนใหญ่นักเรียนชายจะมีคะแนนเฉลี่ยมากกว่านักเรียนหญิง ส่วนด้านคุณภาพของแบบทดสอบก็พอจะมีความเชื่อมั่นอันเป็นที่มั่นใจได้ แต่ปรากฏว่า ค่าอำนาจจำแนกระหว่าง

เพราะค่ามาก ทั้งนี้อาจเป็นเพราะว่า ความสามารถทั้งนักเรียนชายและหญิงนั้นปาน ๆ กัน นั่นคือนักเรียนชายและหญิงจะใช้แบบทดสอบที่เป็นตัวหนังสือหรือรูปภาพก็ได้ใกล้เคียงกัน.

5. สำหรับสหสัมพันธ์พหุคูณของอำนาจจำแนก (Multiple serial correlation) แยกทำการวิเคราะห์ 2 ลักษณะคือ

ก. สหสัมพันธ์พหุคูณของอำนาจจำแนกระหว่างชั้นเรียน = .921

ข. สหสัมพันธ์พหุคูณของอำนาจจำแนกระหว่างอายุ = .533

6. สมการพยากรณ์ระดับชั้นเรียนและระดับอายุของนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4, และปีที่ 7

ก. สมการพยากรณ์ระดับชั้นเรียนของนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4, และปีที่ 7

$$\bar{V} = .01503 X_1 + .04427 X_2 + .03748 X_3 + .03497 X_4 + 1.79910$$

ข. สมการพยากรณ์ระดับอายุของนักเรียนชั้นประถมศึกษาปีที่ 2, ปีที่ 4, และปีที่ 7

$$\bar{V} = .00577 X_1 + .02494 X_2 + .02850 X_3 + .02621 X_4 + 2.50777$$

ทั้ง 2 สมการมี

$\bar{V}$	แทน ระดับชั้นหรือระดับอายุในหน่วยของคะแนนมาตรฐาน
X_1	แทน คะแนนที่ได้ในแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ
X_2	แทน คะแนนที่ได้ในแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ
X_3	แทน คะแนนที่ได้ในแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ
X_4	แทน คะแนนที่ได้ในแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

ข้อเสนอแนะ

ผลจากการวิเคราะห์จากแบบทดสอบทั้ง 4 ฉบับ ทำให้ทราบว่า นักเรียนในชั้นประถมศึกษา จะทำแบบทดสอบที่เป็นรูปภาพได้คะแนนดีกว่าแบบทดสอบที่เป็นตัวหนังสือ จึงขอเสนอแนะตามผลการวิเคราะห์ดังต่อไปนี้

1. ในการสร้างแบบทดสอบในระดับชั้นประถมศึกษา ควรสร้างแบบทดสอบประเภทรูปภาพ มากกว่าแบบทดสอบประเภทตัวหนังสือ
2. ในการทดสอบเด็กชั้นประถมศึกษาตามระดับชั้นเรียน ควรทดสอบด้วยแบบทดสอบประเภทรูปภาพ ในระดับชั้นประถมศึกษาปีที่ 1, ปีที่ 2, และปีที่ 3 มากที่สุด ส่วนชั้นประถมศึกษาปีที่ 4 ควรทดสอบด้วยแบบทดสอบประเภทรูปภาพและตัวหนังสือประกอบในการประเมินผลการเรียนในห้องเรียน ชั้นประถมศึกษาปีที่ 5, ปีที่ 6, และปีที่ 7 ควรทดสอบด้วยแบบทดสอบประเภทตัวหนังสือ แต่อย่างใดก็ได้ผลเที่ยงตรง
3. ในการทดสอบเด็กชั้นประถมศึกษาตามระดับอายุ ควรทดสอบด้วยแบบทดสอบประเภทรูปภาพ ในเด็กอายุตั้งแต่ 7 ปี ถึง 11 ปี ส่วนอายุ 12 ปี ควรทดสอบด้วยแบบทดสอบทั้งประเภทรูปภาพและตัวหนังสือ แต่ในระดับเด็กอายุ 13 ปีขึ้นไป ควรทดสอบด้วยแบบทดสอบที่เป็นตัวหนังสือ แต่อย่างใดก็ได้
4. ในการทดสอบเด็กชั้นประถมศึกษาตามความแตกต่างระหว่างเพศ ควรทดสอบด้วยแบบทดสอบประเภทตัวหนังสือหรือประเภทรูปภาพอย่างใดอย่างหนึ่งก็ได้ผลเช่นกัน
5. ในการเขียนแบบเรียนและทำอุปกรณ์การสอนบางประเภทในเด็กชั้นประถมศึกษาปีที่ 1 ถึงชั้นประถมศึกษาปีที่ 4 ควรให้มีรูปภาพประกอบให้มาก
6. ในการวิจัยต่อไป ควรมีการวิจัยเปรียบเทียบระหว่างแบบทดสอบที่เป็นรูปภาพและตัวหนังสือเกี่ยวกับคำกริยาต่าง ๆ ว่า จะได้ผลเหมือนกับการวิจัยครั้งนี้หรือไม่



บรรณานุกรม

กองการประถมศึกษา, กรมสามัญศึกษา รายงานประจำปี 2504 โรงพิมพ์โรงเรียนการช่าง

วุฒศึกษา (แผนกการพิมพ์) 2506, 229 หน้า.

ชวาล แพทย์กุล เทคนิคการวัดผล โรงพิมพ์วัฒนาพานิช 2509, 452 หน้า.

ศึกษาธิการ, กระทรวง หลักสูตรประโยคประถมศึกษาตอนต้น โรงพิมพ์โรงเรียนการช่างวุฒศึกษา
(แผนกการพิมพ์) 2503, 59 หน้า.

Ahmann, J. Stanley and Glock, Marvin D., Evaluation pupil growth, Boston, Allyn and Bacon, 1958, 605 pp.

Anastasi, Anne, Psychological Testing, The Mcmillan Company, New York, 1961, 675 pp.

Andreas, Burton G., Experimental Psychology, John Wiley & Sons, Inc., New York, 1960, 595 pp.

Boger, Jack H., "An Experimental Study of TheEffects of Perceptual Training on Group I.Q. Test Score of Elementary Pupils in Rural Ungraged School," Journal of Educational Research, 46 : 43 - 52, September, 1952.

Darcy, Natalie T., "The Performance of Bilingual Puerto RicanChildren on Verbal And on Nonlanguage Test of Intelligence," Journal of Educational Research, 45 : 499 - 506, March, 1952.

Deese, James, The Psychology of Learning, McGraw Hill Book Company, Inc, New York, 1960, 478 pp.

Edmonds, William S., "Sex Difference in The Verbal Ability Socio-Economically Depressed Group," The Journal of Educational Research, 58 : 61 - 64, October, 1964.

Edwards, Allen L., Experimental Design in Psychological Research, Rinehart & Company, Inc., New York, 1958, 446 pp.

Fan, Chung-Teh, Item Analysis Table, Educational Testing Service, Princeton, New Jersey, 1952, 32 pp.

Flanagan, John C., Flanagan Aptitude Classification Test, Science Research Associates, Inc., Chicago, 1958, pp 20 - 24.

Garrett, Henry E., Statistics in Psychology and Education, Vakils, Feffer and Simons Private Ltd., Bombay, 1966, 491 pp.

Gendersen, Richard D., and Feldt, Leonard S., "The Relationship of Difference Between Verbal and Nonverbal Intelligence Score to Achievement," Journal of Educational Psychology, 51 : 115 - 121, 1960.

- Greene, Edward B., Measurements of human behavior, The Odyssey Press, New York, 1952, 790 pp.
- Greene, Harry A., Iorgensen, Albert N., and Gerberich, J. Raymon, Measurement and evaluation in the elementary school, Longmans, Green And Co., New York, London, Toronto, 1953, 617 pp.
- Guilford, J.P., Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, New York, 1950, 633 pp.
- Gulliksen, Harold, Theory of Mental Test, John Wiley & Sons, Inc., New York, 1967, 486 pp.
- Hage, Deen S., and Stroud, James B., "Reading Proficiency and Intelligence Scores, Verbal and Nonverbal," Journal of Educational Research, 52 : 258 - 262, March 1959.
- Jersild, Arthur T., Child Psychology, Maruzen Company, Ltd, 1961, 506 pp.
- Killell, Jack E., "Bilingualism And Language Nonlanguage Intelligence Score of Third-Grade Children," Journal of Educational Research, 52 : 263 - 268, March, 1959.
- Mcneamar, Quinn, Psychological Statistic, Hohn Wiley & Sons, Inc., New York, 1962, 451 pp.
- Piaget, Jean, The Language and Thought of The Child, Routledge and Kegan Paul, Ltd, London, 1952, 243 pp.
- Rankin, Paul T., "Listening Ability," Preceeding of The Ohio State Educational Conferenced, Ohio State University, 1929, pp. 172 - 183.
- Ross, C.C., Measurement in Today's School, Prentice - Hall, Inc., New York, 1954, 485 pp.
- Spiegel, Murray R., Theory and Problems of Statistic, Schaum Publishing Co., New York, 1961, 359 pp.
- Wert, James E., Neidt, Charles O., and Ahmann, J. Stanley, Statistical Method in Education and Psychological Research, Appleton - Century-Crofts, New York, 1954, 435 pp.





1. คะแนนเฉลี่ย

$$\bar{X} = \frac{\sum X}{N}$$

2. คะแนนความเบี่ยงเบนมาตรฐาน (standard deviation)

$$s = \sqrt{\frac{N \sum X^2 - (\sum X)^2}{N(N-1)}}$$

3. ค่าสหสัมพันธ์ (correlation)

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

4. ค่าความเชื่อมั่น (reliability)

$$r_{tt} = \frac{n_t^2 - M(n - M)}{\sigma_t^2 (n-1)}$$

5. ความคลาดเคลื่อนมาตรฐานของคะแนน (standard error of measurement)

$$SE_{meas.} = \sigma_1 \sqrt{1 - r_{tt}}$$

6. สหสัมพันธ์ของอำนาจจำแนก (Biserial and serial correlation)

ก. Biserial correlation

$$r_p = \frac{d}{\sigma} \sqrt{pq}$$

ข. serial correlation

$$r_p = \frac{\sum [(y_l - y_h) \bar{x}]}{\sigma \sqrt{\frac{\sum (y_l - y_h)^2}{p}}}$$

7. ทดสอบนัยสำคัญของสหสัมพันธ์อำนาจจำแนก

$$t = \frac{\sqrt{r^2(N-2)}}{1-r^2}$$

8. ความคลาดเคลื่อนมาตรฐานของสหสัมพันธ์

$$\sigma_r = \frac{1-r^2}{\sqrt{N}}$$

9. Simultaneous equation of discriminant analysis

$$\sum x_1 y = a_1 \sum x_1^2 + a_2 \sum x_1 x_2 + a_3 \sum x_1 x_3 + a_4 \sum x_1 x_4$$

$$\sum x_2 y = a_1 \sum x_1 x_2 + a_2 \sum x_2^2 + a_3 \sum x_2 x_3 + a_4 \sum x_2 x_4$$

$$\sum x_3 y = a_1 \sum x_1 x_3 + a_2 \sum x_2 x_3 + a_3 \sum x_3^2 + a_4 \sum x_3 x_4$$

$$\sum x_4 y = a_1 \sum x_1 x_4 + a_2 \sum x_2 x_4 + a_3 \sum x_3 x_4 + a_4 \sum x_4^2$$

10. Multiple serial correlation

$$R_p = \sqrt{\frac{\Delta}{N \sum \frac{(y_1 - y_h)^2}{p}}}$$

11. ทดสอบนัยสำคัญของ R_p

$$F_{m, N-m-1} = \frac{\Delta (N-m-1)}{[N \sum \frac{(y_1 - y_h)^2}{p} - \Delta] m}$$

12. ทดสอบนัยสำคัญของความแปรปรวน

$$F = \frac{s_1^2}{s_2^2}$$

13. ทดสอบนัยสำคัญของความแปรปรวนควาย Analysis of variance

$$F = \frac{SS_{\text{between}}}{SS_{\text{within}}}$$

14. ทดสอบนัยสำคัญของความแตกต่างของคะแนนเฉลี่ย

$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\text{diff.}}}$$

$$SE_{\text{diff.}} = \sqrt{s_{m_1}^2 + s_{m_2}^2}$$

$$SE_{\text{diff.}} = \sqrt{s_{m_1}^2 + s_{m_2}^2 - 2r s_{m_1} s_{m_2}}$$



ตาราง 31. ค่า P_H , P_L , p , r และ Δ ที่ได้จากกราฟวิเคราะห์แบบทดสอบ
อุปมาอุปไมย

ข้อที่	P_H	P_L	p	r	Δ	ข้อที่	P_H	P_L	p	r	Δ
1	96	59	81	.55	9.5	21	81	11	45	.69	12.5
2	99	48	80	.73	9.6	22	85	07	44	.76	13.6
3	96	56	79	.57	9.7	23	78	11	43	.67	13.7
4	99	41	77	.76	10.0	24	59	26	42	.34	13.8
5	93	56	77	.49	10.1	25	52	33	42	.20	13.8
6	99	37	75	.78	10.2	26	56	04	41	.80	14.0
7	74	56	65	.20	11.4	27	70	11	38	.61	14.2
8	99	11	62	.87	11.8	28	81	04	38	.78	14.2
9	96	18	61	.78	11.9	29	70	07	35	.66	14.5
10	93	18	58	.74	12.2	30	74	04	34	.74	14.6
11	85	26	57	.59	12.3	31	67	07	34	.65	14.7
12	89	22	57	.67	12.3	32	52	18	34	.37	14.6
13	89	22	57	.67	12.3	33	67	07	34	.65	14.7
14	85	26	57	.59	12.3	34	52	15	32	.41	14.8
15	89	18	55	.69	12.5	35	48	15	32	.34	14.8
16	85	22	54	.62	12.6	36	56	04	26	.64	15.6
17	85	22	54	.62	12.6	37	30	11	20	.28	16.4
18	78	22	50	.55	13.0	38	26	11	18	.23	16.7
19	85	11	47	.72	13.3	39	37	04	18	.52	16.7
20	81	11	45	.69	13.5	40	26	04	13	.42	17.4

ตาราง 32. ค่า P_H , P_L , p , r และ Δ ที่ไดจากการวิเคราะห์แบบทดสอบ
ไม่เข้าพวก

ข้อที่	P_H	P_L	p	r	Δ	ข้อที่	P_H	P_L	p	r	Δ
1	99	48	80	.73	9.6	21	78	15	46	.62	13.4
2	99	48	80	.73	9.6	22	63	26	44	.38	13.6
3	99	32	74	.79	10.5	23	59	30	44	.30	13.6
4	99	30	72	.80	10.6	24	70	18	43	.52	13.7
5	96	37	71	.60	10.8	25	63	22	42	.42	13.8
6	96	30	68	.72	11.2	26	56	26	41	.31	13.9
7	99	18	66	.85	11.3	27	48	30	39	.20	14.1
8	96	26	66	.74	11.4	28	63	15	38	.50	14.3
9	96	22	64	.76	11.6	29	70	11	38	.61	14.2
10	85	41	64	.47	11.5	30	56	22	38	.36	14.2
11	89	33	63	.59	11.7	31	67	11	37	.59	14.3
12	96	18	61	.70	11.9	32	67	11	37	.59	14.3
13	96	18	61	.70	11.9	33	48	22	35	.29	14.6
14	96	18	61	.70	11.9	34	52	15	32	.41	14.8
15	93	22	60	.71	11.9	35	48	15	30	.33	15.1
16	93	18	58	.74	12.2	36	44	18	30	.30	15.1
17	78	30	54	.48	12.5	37	44	15	29	.34	15.3
18	89	11	50	.75	13.0	38	52	07	27	.55	15.5
19	78	18	48	.59	13.2	39	30	15	22	.21	16.1
20	89	07	47	.79	13.3	40	48	01	18	.71	16.6

ตาราง 33. การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกิติ
ของแบบทดสอบบุปผาอุปไมยประเภทตัวหนังสือ

X	f.	cf.	cf + $\frac{1}{2}f.$	cp.	T
40	.	.	.	.	.
39	.	.	.	.	.
38	1	1316	1315.5	99.9109	84
37	1	1315	1314.5	99.9229	82
36	.	1314	1314.0	99.8469	80
35	.	1314	1314.0	99.8469	80
34	6	1314	1311.0	99.8189	77
33	14	1308	1301.0	98.8591	73
32	17	1294	1285.5	97.6813	70
31	25	1277	1264.5	96.0856	68
30	36	1252	1234.0	93.7680	66
29	36	1216	1198.0	91.0324	64
28	38	1180	1161.0	88.2209	62
27	40	1142	1122.0	85.2574	61
26	39	1102	1092.5	83.0158	60
25	45	1063	1040.5	79.0645	58
24	40	1018	998.0	75.8350	57
23	30	978	963.0	73.1755	56
22	42	948	927.0	70.4399	55
21	58	906	877.0	66.6406	54
20	48	848	824.0	62.6133	53
19	48	800	776.0	58.9659	52
18	55	752	724.5	55.0526	51
17	67	697	663.5	50.1737	50
16	52	630	604.0	45.8961	49
15	61	578	547.5	41.6029	48
14	67	517	483.5	36.7397	47
13	83	450	408.5	31.0407	45
12	74	367	330.0	25.0757	43
11	69	293	258.5	19.6426	41
10	51	224	198.5	15.0834	40
9	53	173	146.5	11.1321	38
8	49	120	95.5	7.2568	36
7	21	71	60.5	4.5972	33
6	19	50	40.5	3.0775	32
5	14	31	24.0	1.8237	29
4	8	17	13.0	.9878	27
3	6	9	6.0	.4559	24
2	3	3	1.5	.1139	20

ตาราง 34. การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกติ
ของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ

X	f.	cf.	cf + $\frac{1}{2}f.$	cp.	T
40	.	.	.	.	.
39	.	.	.	.	.
38	.	.	.	.	.
37	2	1316	1315.0	99.9229	82
36	7	1314	1310.5	99.5810	76
35	9	1307	1302.5	98.9731	73
34	19	1298	1288.5	97.9092	70
33	17	1279	1270.5	96.5415	68
32	28	1262	1248.0	94.8318	66
31	32	1234	1218.0	92.5522	64
30	44	1202	1180.0	89.6647	62
29	46	1158	1135.0	86.2452	61
28	54	112	1085.0	82.4459	59
27	37	1058	1039.5	78.9885	58
26	47	1021	997.5	75.7970	57
25	45	974	951.5	72.3016	56
24	36	929	901.0	69.2242	55
23	30	893	878.0	66.7166	54
22	49	863	838.5	63.7151	53
21	62	814	783.0	59.4978	52
20	80	752	712.0	54.1027	51
19	74	672	635.0	48.2517	49
18	86	598	555.0	42.1728	48
17	69	512	477.5	36.2538	46
16	84	443	401.0	30.4708	45
15	73	359	322.5	25.2657	43
14	76	286	248.0	18.8447	41
13	52	210	184.0	13.9816	39
12	40	158	138.0	10.4862	38
11	41	118	97.5	7.4087	36
10	30	77	62.0	4.7111	33
9	10	47	42.5	3.2294	32
8	17	37	28.5	2.1656	30
7	8	20	16.0	1.2158	28
6	10	12	7.0	.5319	16
5	1	2	1.5	.1139	20
4	1	1	.5	.0379	16
3	.	.	.	.	.
2	.	.	.	.	.
1	.	.	.	.	.

ตาราง 35. การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ประกิติ
ของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ

X	f.	cf.	cf + $\frac{1}{2}f.$	cp.	T
40	.	.	.	.	.
39	.	.	.	.	.
38	.	.	.	.	.
37	.	.	.	.	.
36	.	.	.	.	.
35	1	1315	1315.5	99.9609	84
34	5	1315	1312.5	99.7329	78
33	11	1310	1304.5	99.1250	74
32	17	1299	1290.5	98.0612	71
31	19	1282	1271.5	96.6175	68
30	26	1263	1250.0	94.9838	65
29	29	1237	1222.5	92.8941	64
28	49	1208	1183.5	89.9306	63
27	46	1159	1136.0	86.3212	61
26	45	1113	1090.5	82.8638	59
25	53	1068	1041.5	79.1405	58
24	46	1015	992.0	75.3791	57
23	49	969	944.5	71.7697	56
22	52	920	894.0	67.9324	55
21	45	868	845.5	64.2470	54
20	56	823	795.0	60.4097	53
19	52	767	741.6	56.3064	52
18	37	715	696.5	52.9249	51
17	71	678	642.5	48.8216	50
16	54	607	580.0	44.0725	48
15	75	553	515.5	39.1713	47
14	60	498	448.0	34.0422	46
13	76	418	380.0	28.8751	44
12	73	342	305.5	23.4420	43
11	55	269	241.5	18.3509	41
10	48	214	190.0	14.4375	40
9	42	166	145.0	11.0181	38
8	53	124	97.5	7.4087	36
7	31	71	55.5	4.2173	33
6	16	40	32.0	2.4316	30
5	15	24	16.5	1.2538	28
4	6	9	6.0	.4559	24
3	2	3	2.0	.1520	20
2	.	1	1.0	.0760	18
1	1	1	.5	.0380	17

ตาราง 36. การแจกแจงความถี่, เปอร์เซนไทล์, และคะแนน T - ชรักติ
ของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ

X	f.	cf.	cf + $\frac{1}{2}f.$	op.	T
40	.	.	.	.	.
39	.	.	.	.	.
38	2	1315	1315.0	99.9229	82
37	1	1314	1313.5	99.8089	79
36	4	1313	1311.0	99.6190	77
35	4	1309	1307.0	99.3150	75
34	5	1305	1302.5	98.9731	73
33	28	1300	1286.0	97.7193	70
32	28	1272	1258.0	95.5916	67
31	41	1244	1223.5	92.9701	65
30	62	1203	1172.0	89.0568	63
29	56	1141	1113.0	84.5735	60
28	77	1085	1046.5	89.5204	58
27	84	1008	966.0	73.4034	56
26	92	924	878.0	66.7166	54
25	97	832	783.5	59.5358	52
24	94	735	688.0	52.2791	51
23	74	641	604.0	45.8961	49
22	59	567	637.5	40.8430	48
21	72	508	472.0	35.8659	46
20	65	436	403.5	30.6608	45
19	46	371	348.0	26.4435	43
18	77	325	286.5	21.7703	42
17	45	248	225.5	17.1351	40
16	40	203	183.0	13.9056	39
15	25	163	160.5	12.1959	38
14	33	138	121.5	9.2324	37
13	35	105	87.5	6.6489	35
12	19	770	60.5	4.5972	33
11	17	51	42.5	3.2294	32
10	17	34	25.5	1.9377	29
9	6	17	14.0	1.0638	27
8	10	11	6.0	.4559	24
7	.	1	1.0	.0760	19
6	.	1	1.0	.0760	19
5	.	1	1.0	.0760	19
4	.	1	1.0	.0760	19
3	1	1	.5	.0383	17
2	.	.	.	.	.
1	.	.	.	.	.

ตาราง 37. ค่า N , $\sum X$, $\sum X^2$ ของแบบทดสอบทั้ง 4 ฉบับ จำแนกตามระดับชั้นเรียน

ระดับชั้นเรียน	N	$\sum x_1$	$\sum x_1^2$	$\sum x_2$	$\sum x_2^2$	$\sum x_3$	$\sum x_3^2$	$\sum x_4$	$\sum x_4^2$
ชั้นประถมศึกษาปีที่ 2	424	5037	66909	6114	94614	5059	67419	7718	151254
ชั้นประถมศึกษาปีที่ 4	467	7823	146153	9250	195494	8109	155855	10825	261845
ชั้นประถมศึกษาปีที่ 7	425	10577	275997	11306	312328	10563	272419	11587	322937
รวมทั้งหมด	1316	23437	489059	26670	602436	23731	495693	30130	736036

ตาราง 38. ค่า N , $\sum X$, $\sum X^2$ ของแบบทดสอบทั้ง 4 ฉบับ จำแนกตามอายุ

ระดับอายุ	N	$\sum x_1$	$\sum x_1^2$	$\sum x_2$	$\sum x_2^2$	$\sum x_3$	$\sum x_3^2$	$\sum x_4$	$\sum x_4^2$
7 ปี	74	868	11602	1072	16806	846	10974	1235	22829
8 ปี	154	1891	25735	2268	35674	1879	25687	2827	55319
9 ปี	157	2173	34905	2639	49487	2294	38548	3187	69729
10 ปี	208	3297	60395	3916	80720	3427	63565	4664	110326
11 ปี	189	3093	56099	3604	73410	3115	57711	4271	101109
12 ปี	157	3129	71653	3525	86197	3205	73523	3918	102584
13 ปี	211	5019	128063	5351	143444	4974	124584	5564	150206
14 ปี	108	2628	67588	2796	76798	2590	65252	2885	79537
15 ปี	58	1339	33019	1500	40500	1401	35849	1579	44397
รวมทั้งหมด	1316	23437	489059	26670	602436	23731	495693	30130	736036

ตาราง 39. ค่า N , $\sum X$, $\sum X^2$ ของแบบทดสอบ 4 ฉบับ แยกตามความแตกต่างระหว่างเพศ

ระดับชั้นเรียน	เพศ	N	$\sum X_1$	$\sum X_1^2$	$\sum X_2$	$\sum X_2^2$	$\sum X_3$	$\sum X_3^2$	$\sum X_4$	$\sum X_4^2$
ชั้นประถมศึกษาปีที่ 2	ชาย	232	1821	38643	3379	53253	2861	39271	4299	85525
	หญิง	192	2216	28266	2735	41361	2198	28148	3419	65729
ชั้นประถมศึกษาปีที่ 4	ชาย	259	4477	85819	5262	113668	4675	92065	6147	151149
	หญิง	208	3346	60334	3988	81826	3434	63790	4678	110696
ชั้นประถมศึกษาปีที่ 7	ชาย	246	6180	162786	6677	118011	6216	162816	6838	194322
	หญิง	179	4397	113211	4629	124317	4347	109603	4749	128615
รวมทั้งหมด		1316	23437	489059	26670	602436	23731	495693	30130	736036

ตาราง 40. ค่า $\sum XY$ ของแบบทดสอบทั้ง 4 ฉบับ

แบบทดสอบ	X_1	X_2	X_3	X_4
X_1	—	534556	497884	551352
X_2		—	523469	653063
X_3			—	584727
X_4				—

ตาราง 41. ค่า intercorrelation ของแบบทดสอบทั้ง 4 ฉบับ

แบบทดสอบ	X_1	X_2	X_3	X_4
X_1	—	.8942	.9817	.2565
X_2		—	.6566	.7935
X_3			—	.7399
X_4				—

ตาราง 42. ค่า $\sum X$ ในการวิเคราะห์ความแปรปรวนระหว่างชนิดและประเภทของ
แบบทดสอบควย

ชนิด \ ประเภท	ตัวหนังสือ	รูปภาพ	รวม
ดูมาดูไมย	23437	26670	50107
ไม่เขาพวก	23731	30130	53861
รวม	47168	56800	103968

ตาราง 47. ค่า t ในการทดสอบนัยสำคัญของคะแนนเฉลี่ยระหว่างอายุของ
แบบทดสอบไมเชาพวกประเภทรูปภาพ

ระดับอายุ	7	8	9	10	11	12	13	14	15
7 ปี	—	2.2829**	4.5995**	7.7747**	8.0434**	10.6112**	13.8411**	12.6867**	15.2945**
8 ปี		—	3.3873**	8.0260**	8.4394**	14.4446**	17.8335**	14.3943**	20.4875**
9 ปี			—	3.6477**	3.9725**	7.3438**	11.3867**	9.9051**	13.3372**
10 ปี				—	.3411	4.4101**	8.5721**	7.2808**	10.8057**
11 ปี					—	4.1316**	8.2714**	7.0259**	10.5260**
12 ปี						—	2.6969**	2.7453**	4.4466**
13 ปี							—	.6350	2.2627**
14 ปี								—	.9698
15 ปี									—

ตาราง 48. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ
ตามรายชั้นเรียน

ระดับชั้นเรียน	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
ป. 2	424	0.3222		-0.35863	0.39917900	11.8797
ป. 4	467	0.3549	0.35863	0.00032	0.00000028	16.7516
ป. 7	425	0.3229	0.35895	0.35895	0.39902478	24.8871
Total	1316	1.0000			0.7982046	
$r_{p.ser}$			7.3822			
correct $r_{p.ser}$			.70930			
			.74973			

ตาราง 49. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบฉบับใหม่ประเภทภาพ
ตามรายชั้นเรียน

ระดับชั้นเรียน	N	p	y	$y_1 - y_h$	$(y_1 - y_h)^2/p$	$\bar{x}$
ป. 2	424	0.3222		-0.35863	0.39918179	14.4198
ป. 4	467	0.3549	0.35863	0.00032	0.00000028	19.8073
ป. 7	425	0.3229	0.35895	0.35895	0.39902478	26.6024
Total	1316	1.0000			0.7982046	
6	6.86330					
$r_{p.ser}$	.71496					
correct $r_{p.ser}$	.75576					

ตาราง 50. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกที่เป็นตัวหนังสือ
ตามรายชั้นเรียน

ระดับชั้นเรียน	N	p	y	$y_1 - y_h$	$(y_1 - y_h)^2/p$	$\bar{x}$
ป. 2	424	0.3222		-0.35863	0.39918179	11.9316
ป. 4	467	0.3549	0.35863	0.00032	0.00000028	17.3640
ป. 7	425	0.3229	0.35895	0.35895	0.39902478	24.8541
Total	1316	1.0000			0.7982046	
6	7.17831					
$r_{p.ser}$	.72473					
correct $r_{p.ser}$	.76459					

ตาราง 51. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกที่เป็นรูปภาพ
ตามรายชั้นเรียน

ระดับชั้นเรียน	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
ป. 2	424	0.3222	0.35863	-0.35863	0.3991817900	18.2028
ป. 4	467	0.3549	0.35895	0.00032	0.0000000028	23.1799
ป. 7	425	0.3229		0.35895	0.3990247800	27.2635
Total	1316	1.0000			0.7982046	
6	5.92770					
r _{p.ser}	.61662					
correct p.ser	.65978					

ตาราง 52. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทตัวหนังสือ
ตามรายอายุ

อายุ	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
7	74	0.05623		-0.11320	0.22834	11.7297
8	154	0.11702	0.11320	-0.14293	0.17458	12.2792
9	157	0.11930	0.25613	-0.08758	0.06429	13.8408
10	208	0.15805	0.34371	-0.05217	0.01720	15.8510
13	211	0.16033	0.39588	0.02147	0.00093	23.7867
11	189	0.14361	0.38341	0.06875	0.03293	16.3651
12	157	0.11930	0.31466	0.10765	0.09715	19.9299
14	108	0.08206	0.20701	0.15781	0.15781	24.7963
15	58	0.04407	0.09321	0.09321	0.19718	23.0862
Total	1316	1.00000			0.97041	
๘	7.3822					
$r_{p.ser}$	.4701					
correct $r_{p.ser}$	.4781					

ตาราง 53. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ
ตามรายอายุ

อายุ	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
7	74	0.05623	0.11320	-0.11320	0.22834	11.4324
8	154	0.11702	0.25613	-0.14293	0.47458	12.2013
9	157	0.11930	0.34371	-0.08758	0.06429	14.6115
10	208	0.15805	0.39588	-0.05217	0.01720	16.4760
13	211	0.16033	0.38341	0.01247	0.00093	23.5735
11	189	0.14361	0.31466	0.06875	0.03293	16.4815
12	157	0.11930	0.20701	0.10765	0.09715	20.4140
14	108	0.08206	0.09321	0.11380	0.15718	23.9815
15	58	0.04407		0.09321	0.19718	24.1552
Total	1316	1.00000			0.98509	
6	6.8633					
r _{p.ser}	.4736					
correctr _{p.ser}	.4016					

ตาราง 54. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทตัวหนึ่งต่อ

ตามรายอายุ

อายุ	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
7	74	0.05623	0.11320	-0.11320	0.22834	11.4324
8	154	0.11702	0.25613	-0.14293	0.17458	12.2013
9	157	0.11930	0.34371	-0.08758	0.06429	14.6115
10	208	0.15805	0.39588	-0.05217	0.01720	16.4760
13	211	0.16033	0.38341	0.01247	0.00093	23.5735
11	189	0.14361	0.31466	0.06875	0.03293	16.4815
12	157	0.11930	0.20701	0.10765	0.09715	20.4140
14	108	0.08206	0.09321	0.11380	0.15718	23.9815
15	58	0.04407		0.09321	0.19718	24.1552
Total	1316	1.00000			0.98509	
6	7.1783					
$r_{p.ser}$	.4848					
correct $r_{p.ser}$	.4930					

ตาราง 55. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทภาพ
ตามรายอายุ

อายุ	N	p	y	$y_l - y_h$	$(y_l - y_h)^2/p$	$\bar{x}$
7	74	0.05623		-0.11320	0.22834	16.6892
			0.11320			
8	154	0.11702		-0.14293	0.17456	18.3571
			0.25613			
9	157	0.11930		-0.08758	0.06429	20.2994
			0.34371			
10	208	0.15805		-0.05217	0.01720	22.4231
			0.39528			
13	211	0.16033		0.01247	0.00093	26.3697
			0.33341			
11	189	0.14361		0.06875	0.03293	22.5979
			0.31466			
12	157	0.11930		0.10765	0.09715	24.9554
			0.20701			
14	108	0.08206		0.11380	0.15781	26.7130
			0.09321			
15	58	0.04407		0.09321	0.19718	27.2241
Total	1316	1.00000			0.97041	
6	5.9277					
$r_{p.ser}$	.4599					
correct $r_{p.ser}$	.4678					

ตาราง 56. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภท
ตัวหนังสือ จำแนกตามเพศ

เพศ	N	percent	Σx	$\bar{x}$
ชาย	737	56.003	13478	18.28765
หญิง	579	43.996	9959	17.20034
รวม	1316	100.000		
r	7.3822			
$r_{p.ser}$	.0731			
correct $r_{p.ser}$	.0915			

ตาราง 57. สหสัมพันธ์อำนาจจำแนกของแบบทดสอบอุปมาอุปไมยประเภทรูปภาพ
จำแนกตามเพศ

เพศ	N	percent	Σx	$\bar{x}$
ชาย	737	56.003	15318	20.78426
หญิง	579	43.996	11352	19.60621
รวม	1316	100.000		
r	6.8633			
$r_{p.ser}$	.0852			
correct $r_{p.ser}$	.1066			

ตาราง 58. สหสัมพันธ์ของอำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทตัวหนังสือ
จำแนกตามเพศ

เพศ	N	percent	Σx	$\bar{x}$
ชาย	737	56.003	13752	18.65943
หญิง	579	43.996	9979	17.23488
รวม	1316	100.000		
ϕ	7.1783			
$r_{p.ser}$	.0985			
correct $r_{p.ser}$	.1232			

ตาราง 59. สหสัมพันธ์อำนาจจำแนกของแบบทดสอบไม่เข้าพวกประเภทรูปภาพ
จำแนกตามเพศ

เพศ	N	percent	Σx	$\bar{x}$
ชาย	737	56.003	17284	23.45183
หญิง	579	43.996	12850	22.19343
รวม	1316	100.000		
ϕ	5.9277			
$r_{p.ser}$	.1054			
correct $r_{p.ser}$	.1318			