

การวิเคราะห์แบบทดสอบ

ด้วยโปรแกรม TAP (Test Analysis Program)

Step by Step



ประเสริฐ สุภีรักษ์

กลุ่มงานวัดและประเมินผลการศึกษา

สพท.นครสวรรค์ 1

คำนำ

การวิเคราะห์แบบทดสอบ ด้วยโปรแกรม TAP : Test Analysis Program (Step by Step) เล่มนี้ ผู้เขียนมีความตั้งใจที่ต้องการให้ผู้ศึกษาสามารถนำขั้นตอนในการวิเคราะห์ TAP ไปใช้ในงาน พัฒนาการเรียนการสอนในชั้นเรียน เพื่อวิเคราะห์การทดสอบของนักเรียนและการวิเคราะห์หาคุณภาพ แบบทดสอบในการแก้ไขปรับปรุงให้มีประสิทธิภาพมากขึ้น จากการที่ผู้เขียนรับผิดชอบงานวัดและ ประเมินผลการศึกษาของเขตพื้นที่การศึกษา ทำให้ทราบปัญหาของครูผู้สอนในการสร้างและพัฒนา เครื่องมือวัดจะมีความยากในขั้นตอนวิเคราะห์ข้อมูล ซึ่งอาจจะต้องพึ่งพาผู้วิเคราะห์ข้อมูลให้ และมักจะ พบว่าข้อมูลที่วิเคราะห์นั้นวิเคราะห์ถูกต้องหรือไม่เพียงไร ตำราเล่มนี้จะช่วยให้สามารถวิเคราะห์ ได้ด้วยตนเองอย่างถูกต้องแม่นยำยิ่งขึ้น เท่ากับเป็นการแก้ไขปัญหาดังกล่าวได้เป็นอย่างดีในระดับหนึ่ง ผู้เขียนขอขอบคุณ นางสาวณนุชกุล สุภีรักษ์ ที่ได้ช่วยแปลเอกสารบางส่วน จากการค้นคว้าของผู้เขียน สุดท้ายผู้เขียนหวังเป็นอย่างยิ่งว่า การนำเอกสารเล่มนี้ไปใช้คงอำนวยความสะดวกต่อผู้ใช้ พอสมควร หากมีข้อบกพร่องประการใดขอน้อมรับคำแนะนำเพื่อการแก้ไขปรับปรุงโอกาสต่อไป

ประเสริฐ สุภีรักษ์

e-mail : supirak5@hotmail.com

สารบัญ

บทที่ 1 บทนำ	1
ความรู้เบื้องต้นเกี่ยวกับ TAP	1
การแสดงผลการวิเคราะห์แบบทดสอบ (TAP output provides)	1
บทที่ 2 การติดตั้งโปรแกรม	3
การเข้าสู่โปรแกรม TAP	3
ขั้นตอนการป้อนข้อมูล	4
ในกรอบหน้าต่าง INPUT ให้เลือก	4
หน้าต่าง Data Editor ป้อนข้อมูลรายการต่างๆ ลงในช่วงว่าง	5
บทที่ 3 ขั้นตอนการวิเคราะห์ข้อมูล	6
ผลการวิเคราะห์ข้อมูล	6
บทที่ 4 ผลลัพธ์ที่ได้จากการวิเคราะห์	9
ส่วนที่ 1. ผลการวิเคราะห์ผู้สอบ (Examinee Analysis)	9
ส่วนที่ 2 ผลการวิเคราะห์ข้อสอบรายข้อและทั้งฉบับ (Item and Test Analysis)	11
ส่วนที่ 3 ผลการวิเคราะห์รายตัวเลือก (Options Analysis)	14
ส่วนที่ 4 การแสดงด้วยแผนภาพ (Graph : Bar Chart and Histogram)	14
บทที่ 5 การเลือกใช้ เมนู Option	16
การกำหนดค่าเปอร์เซ็นต์ของเกรด (Set Percentages If Grades)	16
การกำหนดค่าเปอร์เซ็นต์แบ่งกลุ่มสูงกลุ่มค่า(Set Percentages of Item Discrimination)	18
บทที่ 6 การเลือกใช้ เมนู Analysis	19
การกำหนดจำนวนข้อสอบเพิ่มเติมตามค่าความเชื่อมั่นที่ต้องการ	19
ผลการวิเคราะห์	21
บทที่ 7 การตรวจสอบคุณภาพเครื่องมือวัดผลการศึกษา	22
การพิจารณาค่าความยาก(p) และอำนาจจำแนก(r, D) แบบทดสอบ	35

บทที่ 1

บทนำ

1.1 ความรู้เบื้องต้นเกี่ยวกับ TAP

โปรแกรม TAP (Test Analysis Program) ถูกออกแบบมาอย่างมีประสิทธิภาพ ง่ายต่อการใช้ และเป็นโปรแกรม Freeware สามารถดาวน์โหลดได้จากเว็บไซต์ <http://oak.cats.ohio.edu/~brooksg/tap.htm> พัฒนาโดย บรู๊ก (Brooks, Gordon P.) เป็นชุดโปรแกรมให้ใช้ฟรี เพื่อวิเคราะห์การทดสอบสำหรับครูผู้สอนในระดับโรงเรียนตลอดจนในระดับมหาวิทยาลัย จุดเด่นเป็นลักษณะของ Software เครื่องมือที่อัจฉริยะสำหรับทดสอบออกแบบหลักสูตร ผู้ใช้สามารถป้อนข้อมูลคะแนนทดสอบในรูปแบบตัวอักษร หรือข้อมูลการทดสอบโดยตรง ในการวิเคราะห์สถิติของผู้สอบ ประกอบด้วยเปอร์เซ็นต์ในการตอบถูก การตัดเกรด ช่วงความเชื่อมั่นของนักเรียนแต่ละคน สถิติพื้นฐานของกลุ่มผู้สอบ จำนวนข้อที่ตอบถูกแต่ละคน และจำนวนข้อที่ตอบผิด วิเคราะห์สถิติของข้อสอบและแบบทดสอบ ประกอบด้วย การวิเคราะห์ความยาก อำนาจจำแนก พอยท์ไบซีเรียล ตลอดจนวิเคราะห์คุณภาพเป็นรายตัวเลือก และสามารถแบ่งกลุ่มสูงและกลุ่มต่ำได้ตามกำหนด

1.2 การแสดงผลการวิเคราะห์แบบทดสอบ (TAP output provides) 2 ส่วน

1. ผลการวิเคราะห์ผู้ทดสอบ (Examinee Analysis) ประกอบด้วย ร้อยละความถูกต้อง (Percentage Correct) เกรด ช่วงความเชื่อมั่น สำหรับนักเรียนแต่ละคน และผลรวมคะแนน สถิติเชิงพรรณนา สำหรับกลุ่มผู้สอบ TAP จะสร้างรายการคะแนนและคำตอบที่ผิดของนักเรียน (คำตอบที่นักเรียนตอบผิด) ด้วย เช่น รายงานนักเรียนเป็นรายบุคคล บ่งชี้คะแนนตอบสนองต่อข้อสอบที่ถูกต้องไปจนถึงสิ่งที่ผิดพลาด

2. ผลการวิเคราะห์ข้อสอบและการทดสอบ (Item and Test Analysis) ประกอบด้วย ความยาก ดัชนีค่าอำนาจจำแนก (Discrimination index) รายข้อรายตัวเลือก และ สถิติความแปรปรวน KR-20 , KR-21 ค่าเฉลี่ย ค่าความเบี่ยงเบน ถ้าข้อสอบถูกปล่อย TAP สามารถคำนวณจำนวนข้อสอบเพิ่มเติมตามความต้องการที่เพิ่มค่าความเชื่อมั่นได้

การวิเคราะห์ตัวเลือก (Option Analysis) ประกอบด้วย กลุ่มสูง และกลุ่มต่ำ ความยากของข้อสอบ เพื่อตรวจคำตอบ(คำตอบที่ถูกต้อง) และตัวลวง

ในเอกสารคู่มือการใช้ TAP เล่มนี้ จะนำเสนอขั้นตอนการใช้โปรแกรม TAP Version 6.65 ซึ่งได้ปรับปรุงพัฒนาตั้งแต่ Version 4.xx ถึง Version 6.65 ดังนี้

Version 6.6X :

- แก้ไขข้อบกพร่องหลัก ด้วยรายงานเกรดส่วนบุคคล (ที่เคยเป็นปัญหาการจำแนกสำหรับการตอบสนอง / คำตอบข้อสอบแต่ละบุคคล)
- เพิ่มความสามารถเพื่อให้ผู้ใช้งานระบุ Criteria ด้วยตนเองสำหรับข้อสอบที่มีปัญหา
- เพิ่มความสามารถในการจำแนกผลการวิเคราะห์ Quick Item โดยการปรับ Point Biserial.

Version 6.4X :

- ความสามารถในการป้อนข้อมูลตัวอักษร (A,B,C,D) ดังเช่น ข้อมูลตัวเลข (1,2 ,3, 4) สำหรับเฉลย และคำตอบของนักเรียน

Version 6.XX :

- การนำเข้า input files ที่ดีกว่าทุก version
- HELP เพื่อช่วยในเรื่องของการนำเข้าข้อมูล SPSS ด้วยคำสั่งใน Sytax ที่ใช้ใน SPSS
- ความสามารถในการซ่อน และไม่ซ่อน (HIDE หรือ UNHIDE) ชื่อจาก Output (default เป็น hide, ด้วย sasy button เป็น unhide)
- ปุ่มการใช้งานง่ายสำหรับคำสั่งการ Save ไฟล์ Output ทั่วไปทุกชนิด
- การวิเคราะห์ห้อย่างรวดเร็วสำหรับ ตัวเลือก / ตัวลง
- Output ที่ถูกเพิ่มเติม (เช่น class statistics, hard item annotations) ไปสู่ Individual Grade Reports.
- HELP ในหน้าต่างแก้ไขข้อมูลสำหรับ OHOI UWIVERSITY FOUKS ที่ยอมให้นำเข้าข้อมูล scanned TESTSCORE ที่ง่ายกว่า
- HELP ในหน้าต่างแก้ไขข้อมูล นับจำนวน / ines (บรรทัด) ในหน้าต่างข้อมูล

Version 5.XX :

- เกรดติดบางส่วน (Partial credit) สามารถเพิ่มคะแนน (แต่ไม่วิเคราะห์)
- ข้อสอบแบบหลายตัวเลือก สามารถมีค่ามากกว่า 1 คะแนน (แต่ยังคงวิเคราะห์ที่ 1 คะแนน)
- User Grading Categories ปรับเปลี่ยนใช้มากกว่าเดิม และปรับเปลี่ยนใช้ได้ตามต้องการ

Version 4.XX :

- จำนวนผู้ถูกทดสอบ เพิ่มเป็น 9999
- Histograms และ กราฟแท่งสามารถปรับ, Save และ Print
- จำนวนข้อสอบเพิ่มเป็น 100
- Histograms และ bar charts แบบใหม่ สามารถสร้างได้จาก Full Results (ผลลัพธ์ทั้งหมด)
- ตาราง ความถี่ (ด้วย PR) เพิ่มใน Full Results.
- Split – Half Reliability เพิ่มใน Full Results.

บทที่ 2

การติดตั้งโปรแกรม

2.1 การติดตั้งโปรแกรม TAP

ดาวน์โหลดไฟล์ tap.exe จากเว็บไซต์ <http://oak.cats.ohio.edu/~brooksg/tap.htm> หรือ Copy ไฟล์ tap.exe จาก CD ที่ให้มาพร้อมกับเอกสารนี้ ลงใน My Documents หรือ Copy มาวาง หน้า Desktop เป็นไอคอน ดังภาพประกอบ 1



ภาพประกอบ 1

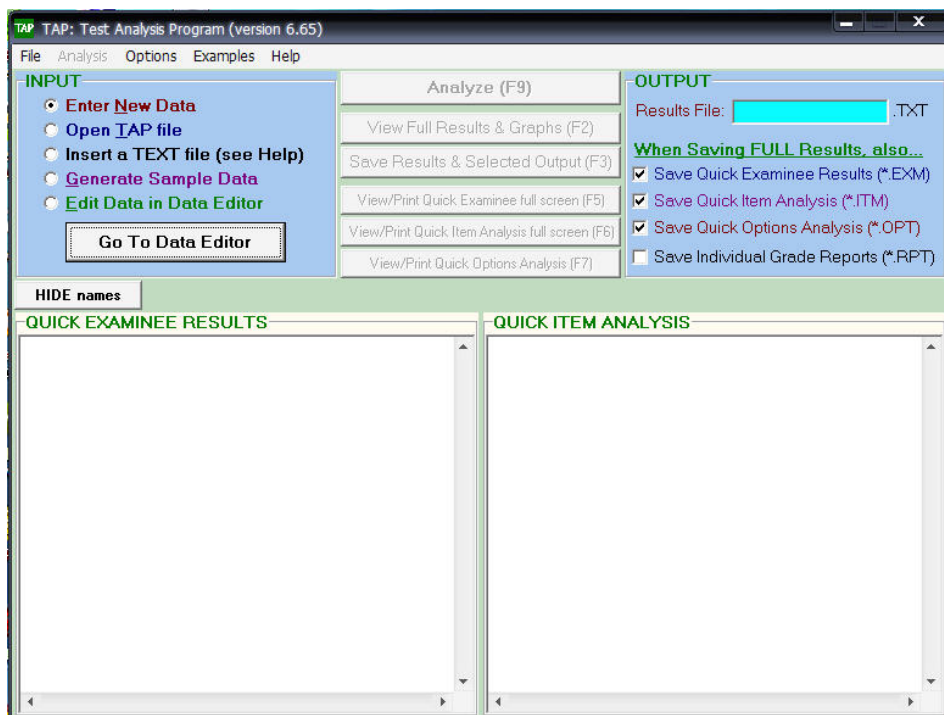
2.2 การเข้าสู่โปรแกรม TAP

ขั้นตอนการเรียกใช้โปรแกรม TAP (Test Analysis Program) for Windows สามารถทำได้โดย



ภาพประกอบ 2

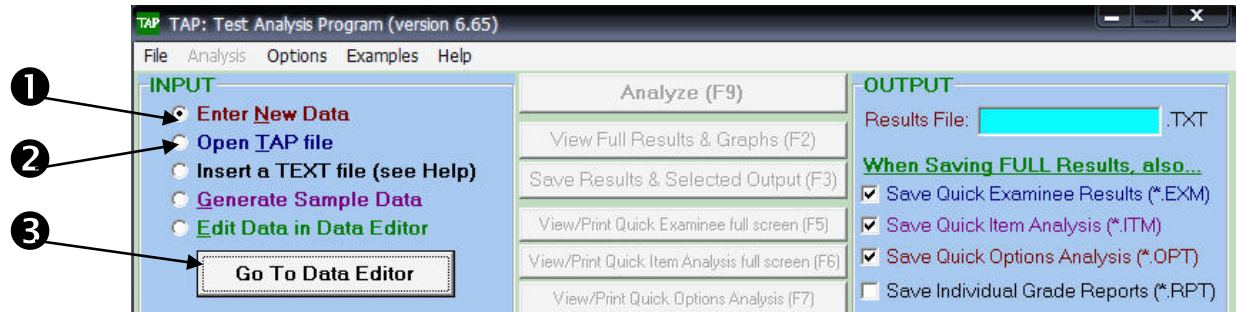
เข้าสู่หน้าจอ TAP : Test Analysis Program (Version 6.65)



ภาพประกอบ 3

2.3 ขั้นตอนการป้อนข้อมูล

2.3.1 ในกรอบหน้าต่าง INPUT ให้เลือก



ภาพประกอบ 4

1. เลือก Enter New Data การสร้าง/ป้อนข้อมูลใหม่
2. หรือ เลือก Open TAP file จากนั้น คลิกที่ปุ่ม GO To Data Editor ในกรณี เปิดข้อมูลเก่าที่สร้างจากโปรแกรม TAP
3. คลิกที่ปุ่ม GO To Data Editor

2.3 .2 หน้าต่าง Data Editor ป้อนข้อมูลรายการต่างๆ ลงในช่วงว่าง ดังนี้

The screenshot shows the TAP Data Editor interface with the following fields and values:

- Title:** MAHT-P5 (70 character maximum)
- Comments:** LAS-2550 (70 character maximum)
- Key:** 12414243233344134213
- Options:** 44444444444444444444
- Include:** vvvvvvvvvvvvvvvvvvvvvvv
- # Students:** 50
- # Test Items:** 20
- Length of Examinee Names/IDs (# characters):** 12

The DATA section contains a table with the following columns: Examinee ID, Item#, and Item#. The table lists 16 rows of student data.

Examinee ID	Item#	Item#
Student035	23414412191344423123	
Student036	24213132433444134343	
Student037	22324323243232321234	
Student038	11441242349243242431	
Student039	12243214444134242	
Student040	11243121411142134	
Student041	14243342344132214	
Student042	12411244323434194214	
Student043	12211143331341134213	
Student044	44114443423424434241	
Student045	41124213231421134213	
Student046	31224144323244411342	
Student047	42112143412314221344	
Student048	41442142133434331113	
Student049	11414249324231332994	
Student050	21432413242134312423	


ภาพประกอบ 5

1. Title ป้อนหัวข้อของแบบทดสอบ หรือวิชาที่จัดสอบ จำนวนตัวอักษรสูงสุด 70 ตัวอักษร
2. Comments ป้อนหมายเหตุ หรืออาจจะระบุว่าแบบทดสอบนี้วัดอะไร จำนวนตัวอักษรสูงสุด 70 ตัวอักษร
3. Key ป้อนเฉลยของข้อสอบ
4. Options ป้อนจำนวนตัวเลือกในข้อสอบแต่ละข้อ
5. Include ระบุว่าข้อใดที่ต้องการวิเคราะห์และข้อใดไม่นำเข้าวิเคราะห์ โดยพิมพ์ Y คือนำข้อสอบข้อนั้นเข้าวิเคราะห์ หรือพิมพ์ N เมื่อไม่นำข้อสอบนั้นเข้ามาวิเคราะห์
6. #Students ป้อนจำนวนนักเรียนหรือผู้สอบทั้งหมด
7. # Test Items ป้อนจำนวนข้อสอบทั้งหมด
8. Length of Examinee Names /IDs (# character) ระบุจำนวนสดมภ์ที่ต้องการป้อนรหัสประจำตัวผู้สอบหรือชื่อผู้สอบ
9. DATA สำหรับป้อนข้อมูลทั้งหมดของผู้สอบ โดยป้อนรหัสประจำตัวผู้สอบตามจำนวนสดมภ์ที่กำหนด และป้อนตัวเลือกแต่ละข้อของผู้สอบแต่ละคน ตามจำนวนที่ระบุ
10. เมื่อป้อนข้อมูลเสร็จสิ้น ให้บันทึกข้อมูลโดยคลิกที่ปุ่ม Save File

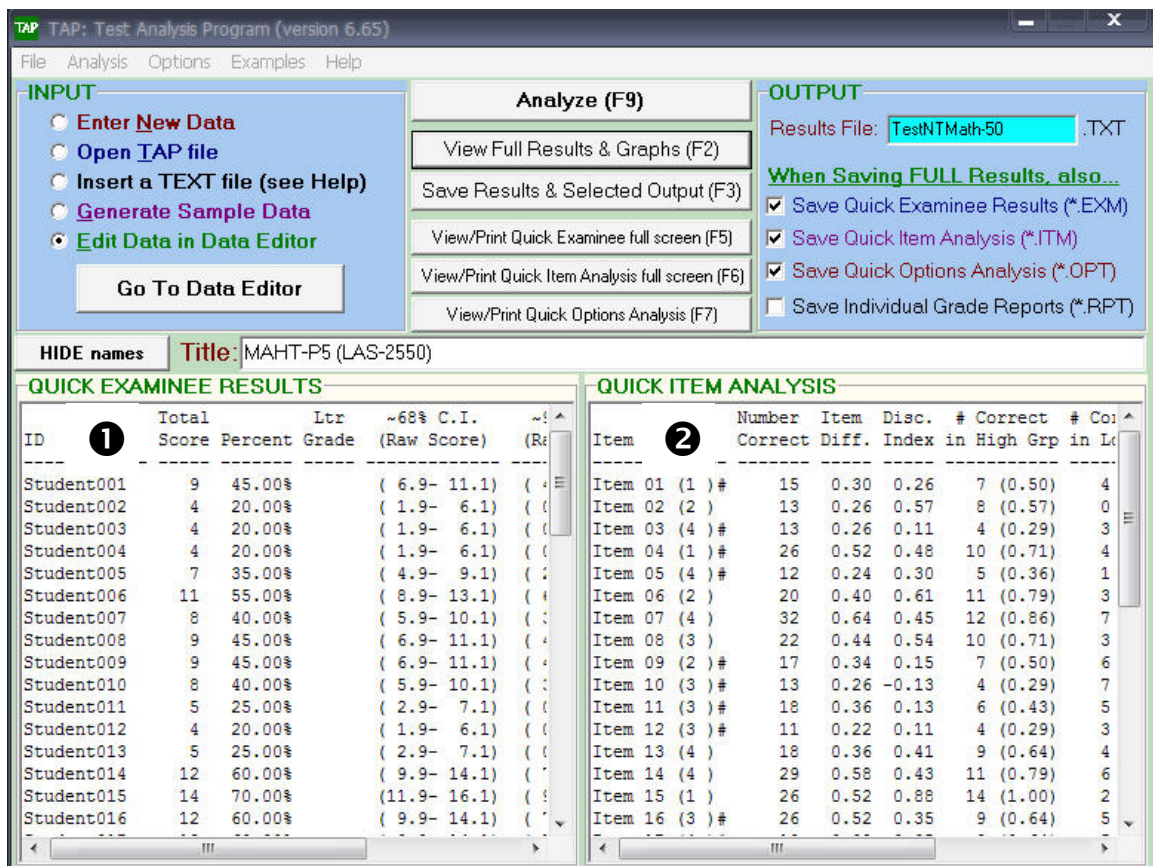
บทที่ 3

ขั้นตอนการวิเคราะห์ข้อมูล

3. ขั้นตอนการวิเคราะห์ข้อมูล

3.1 คลิกที่ปุ่ม  ในหน้าต่าง Data Editor ภาพประกอบ 5 โปรแกรมจะวิเคราะห์ข้อมูล

3.2 ผลการวิเคราะห์ข้อมูล จะแสดงใน 2 หน้าต่างคือหน้าต่าง ดังภาพประกอบ 6



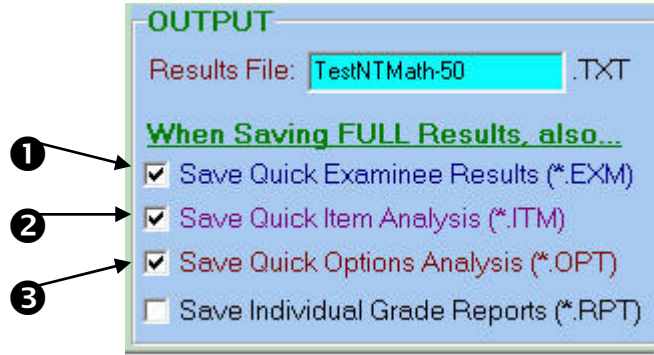
ID	Total Score	Ltr Percent Grade	~68% C.I. (Raw Score)	~1%
Student001	9	45.00%	(6.9- 11.1)	(4.5)
Student002	4	20.00%	(1.9- 6.1)	(1.3)
Student003	4	20.00%	(1.9- 6.1)	(1.3)
Student004	4	20.00%	(1.9- 6.1)	(1.3)
Student005	7	35.00%	(4.9- 9.1)	(2.7)
Student006	11	55.00%	(8.9- 13.1)	(4.5)
Student007	8	40.00%	(5.9- 10.1)	(3.2)
Student008	9	45.00%	(6.9- 11.1)	(3.2)
Student009	9	45.00%	(6.9- 11.1)	(3.2)
Student010	8	40.00%	(5.9- 10.1)	(3.2)
Student011	5	25.00%	(2.9- 7.1)	(1.9)
Student012	4	20.00%	(1.9- 6.1)	(1.3)
Student013	5	25.00%	(2.9- 7.1)	(1.9)
Student014	12	60.00%	(9.9- 14.1)	(4.5)
Student015	14	70.00%	(11.9- 16.1)	(4.5)
Student016	12	60.00%	(9.9- 14.1)	(4.5)

Item	Number Correct	Item Diff.	Item Disc. Index	# Correct in High Grp	# Correct in Ltr
Item 01 (1)#	15	0.30	0.26	7 (0.50)	4
Item 02 (2)	13	0.26	0.57	8 (0.57)	0
Item 03 (4)#	13	0.26	0.11	4 (0.29)	3
Item 04 (1)#	26	0.52	0.48	10 (0.71)	4
Item 05 (4)#	12	0.24	0.30	5 (0.36)	1
Item 06 (2)	20	0.40	0.61	11 (0.79)	3
Item 07 (4)	32	0.64	0.45	12 (0.86)	7
Item 08 (3)	22	0.44	0.54	10 (0.71)	3
Item 09 (2)#	17	0.34	0.15	7 (0.50)	6
Item 10 (3)#	13	0.26	-0.13	4 (0.29)	7
Item 11 (3)#	18	0.36	0.13	6 (0.43)	5
Item 12 (3)#	11	0.22	0.11	4 (0.29)	3
Item 13 (4)	18	0.36	0.41	9 (0.64)	4
Item 14 (4)	29	0.58	0.43	11 (0.79)	6
Item 15 (1)	26	0.52	0.88	14 (1.00)	2
Item 16 (3)#	26	0.52	0.35	9 (0.64)	5

ภาพประกอบ 6

1. หน้าต่าง QUICK EXAMINEE RESULTS แสดงค่าสถิติสำหรับผู้สอบ และ
2. หน้าต่าง QUICK ITEM ANALYSIS แสดงค่าสถิติสำหรับข้อสอบ

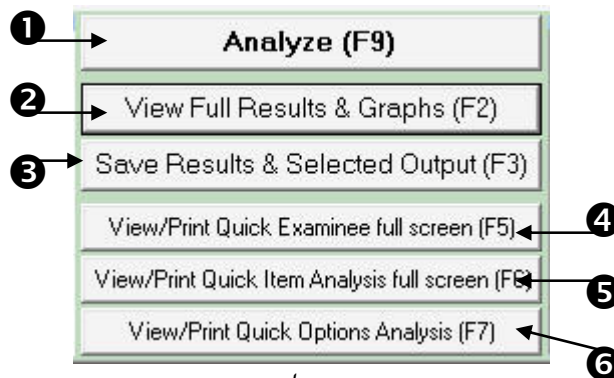
- 3.3 เมื่อต้องการบันทึกผลการวิเคราะห์ทั้งสองหน้าต่างลงไปไฟล์ ให้ไปที่ช่องOUTPUT โดยพิมพ์ชื่อไฟล์ที่ต้องการเก็บผลลัพธ์ในช่อง Result File Name โดยผลลัพธ์จะเก็บไว้ในนามสกุล .Txt และ



ภาพประกอบ 7

1. คลิก Save Quick Examinee Results โปรแกรมจะบันทึกไว้ในชื่อเดียวกับResult File Name แต่นามสกุลต่างกัน คือ .exm
2. คลิก Save Quick Item Analysis โปรแกรมจะบันทึกไว้ในชื่อเดียวกับResult File Name แต่นามสกุลต่างกัน คือ .itm
3. คลิก Save Quick Option Analysis โปรแกรมจะบันทึกไว้ในชื่อเดียวกับResult File Name แต่นามสกุลต่างกัน คือ .opt

- 3.4 การวิเคราะห์ผลอย่างย่อ และต่อเนื่องด้วยการกำหนด Option ใหม่



ภาพประกอบ 8

1. ปุ่ม Analyze (F9) เป็นปุ่มวิเคราะห์ข้อมูลโดยจะแสดงผลอย่างย่อและเมื่อคลิก Option เพื่อกำหนด Option เพิ่มเติมใหม่ เช่น การตัดเกรด หรือร้อยละของกลุ่มสูง- กลุ่มต่ำ สามารถวิเคราะห์อย่างต่อเนื่องได้ ปุ่มAnalyze(F9) จะเปลี่ยนเป็นดังภาพประกอบ 9



ภาพประกอบ 9

2. ปุ่ม View Full Results & Graphs (F2) เป็นปุ่มวิเคราะห์ข้อมูลโดยละเอียด จะให้ผลการวิเคราะห์ผู้สอบและข้อสอบที่มากกว่า นอกจากนี้ยังแสดงแผนภูมิแท่งและฮิสแกรมของคะแนนสอบได้อีกด้วย
3. ปุ่ม Save Results & Selected Output (F3) เป็นปุ่มบันทึกผลการวิเคราะห์โดยละเอียดและเลือกการบันทึกไฟล์ผลลัพธ์
4. ปุ่ม View/Print Quick Examinee Full Screen (F5)) เป็นปุ่มแสดงเฉพาะผลการวิเคราะห์ค่าสถิติของนักเรียนหรือผู้สอบรายบุคคล ทุกรายการ
5. ปุ่ม View/Print Item Analysis Full Screen (F6) เป็นปุ่มแสดงผลการวิเคราะห์ค่าสถิติของแบบทดสอบรายข้อ และคุณภาพ ตลอดจนค่าความเชื่อมั่น(KR-20,KR-21) ของแบบทดสอบทั้งฉบับ ทุกรายการ
6. ปุ่ม View/Print Quick Option Analysis (F7) เป็นปุ่มแสดงผลเฉพาะการวิเคราะห์ค่าสถิติคุณภาพของแบบทดสอบรายข้อ และตัวเลือกตัวลวง (ค่า p ,r (D) ตาม Option ที่กำหนด

บทที่ 4

ผลลัพธ์ที่ได้จากการวิเคราะห์

4. ผลลัพธ์ที่ได้จากการวิเคราะห์โดยละเอียด

ผลลัพธ์ที่ได้จากการวิเคราะห์จะแบ่งออกเป็น 4 ส่วนดังนี้

ส่วนที่ 1. ผลการวิเคราะห์ผู้สอบ (Examinee Analysis) ให้ผลดังนี้

ส่วนที่ 1

ID	Total Score	Percent	Ltr Grade	~68% C.I. (Raw Score)	~95% C.I. (Raw Score)
Student001	9	45.00%	D	(6.9- 11.1)	(4.9- 13.1)
Student002	4	20.00%	F	(1.9- 6.1)	(0.0- 8.1)
Student003	4	20.00%	F	(1.9- 6.1)	(0.0- 8.1)
Student004	4	20.00%	F	(1.9- 6.1)	(0.0- 8.1)
Student005	7	35.00%	F	(4.9- 9.1)	(2.9- 11.1)
Student006	11	55.00%	C	(8.9- 13.1)	(6.9- 15.1)

ภาพประกอบ 10

ID คือรหัสประจำตัวของผู้สอบแต่ละคน

Score คือคะแนนที่ได้ของผู้สอบแต่ละคน

Percent คือสัดส่วนระหว่างคะแนนที่ได้กับคะแนนเต็ม

Ltr Grade คือผลการตัดเกรด ซึ่งผู้วิเคราะห์สามารถกำหนดได้ว่าจะให้ตัดกี่เกรด และเกณฑ์ผ่านคือเท่าไร

68% C.I. คือประมาณค่าคะแนนจริงของผู้สอบแต่ละคน จากสูตร $X \pm SEM$ ที่ระดับความเชื่อมั่น 68%

95% C.I. คือประมาณค่าคะแนนจริงของผู้สอบแต่ละคน จากสูตร $X \pm (1.96)SEM$ ที่ระดับความเชื่อมั่น 95%

ส่วนที่ 1(ต่อ)

Number of Examinees = 50
 Total Possible Score = 20
 Minimum Score = 4.000 = 20.0%
 Maximum Score = 14.000 = 70.0%
 Median Score = 7.000 = 35.0%
 Mean Score = 7.880 = 39.4%
 Standard Deviation = 3.154
 Variance = 9.946
 Skewness = 0.470
 Kurtosis = -1.046

=====
 Frequency Table
 =====

Score	Z Score	Freq.	Rel. Freq.	Cum. Freq.	Cum. Percent	Percentile Rank	Stanine	Normalized Zn Score
4	-1.23	7	0.140	7	14.00	7.0	2	-1.47
5	-0.91	10	0.200	17	34.00	24.0	4	-0.70
6	-0.60	4	0.080	21	42.00	38.0	4	-0.30
7	-0.28	5	0.100	26	52.00	47.0	5	-0.07
8	0.04	4	0.080	30	60.00	56.0	5	0.15
9	0.36	6	0.120	36	72.00	66.0	6	0.41

ภาพประกอบ 11

Number of Examinees คือจำนวนผู้สอบทั้งหมด

Minimum Score คือคะแนนต่ำสุดที่ผู้สอบในกลุ่มทำได้

Maximum Score คือคะแนนสูงสุดที่ผู้สอบในกลุ่มทำได้

Median Score คือคะแนนมัธยฐาน

Mean Score คือคะแนนเฉลี่ย

Standard Deviation คือคะแนนเบี่ยงเบนมาตรฐาน

Variance คือความแปรปรวน

Skewness คือความเบ้

Kurtosis คือความโค้ง

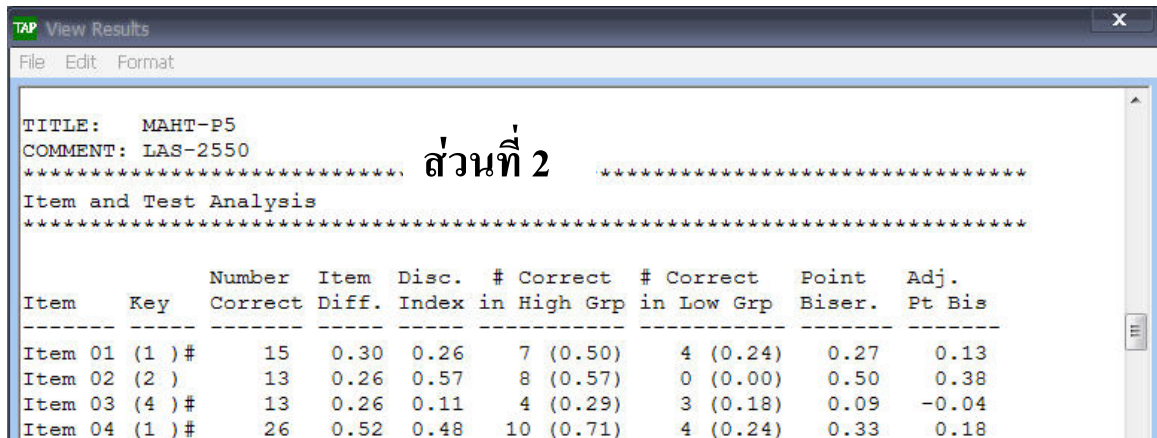
Frequency Table คือ การแสดงแจกแจงความถี่ของคะแนนสอบ

Bar Graph คือแสดงการแจกแจงความถี่ของคะแนนด้วยแผนภูมิแท่ง

Stem-and-Leaf Display คือแสดงการแจกแจงของคะแนนในรูปของแผนภาพต้นไม้

Bar Graph คือแสดงการแจกแจงความถี่ของแต่ละระดับเกรด

ส่วนที่ 2 ผลการวิเคราะห์ข้อสอบรายข้อและทั้งฉบับ (Item and Test Analysis) ให้ผลดังนี้



Item	Key	Number Correct	Item Diff.	Disc. Index	# Correct in High Grp	# Correct in Low Grp	Point Biser.	Adj. Pt Bis
Item 01	(1)#	15	0.30	0.26	7 (0.50)	4 (0.24)	0.27	0.13
Item 02	(2)	13	0.26	0.57	8 (0.57)	0 (0.00)	0.50	0.38
Item 03	(4)#	13	0.26	0.11	4 (0.29)	3 (0.18)	0.09	-0.04
Item 04	(1)#	26	0.52	0.48	10 (0.71)	4 (0.24)	0.33	0.18

ภาพประกอบ 12

Item คือข้อสอบแต่ละข้อ

Number Correct คือจำนวนผู้สอบที่ตอบถูกในข้อนั้น

Item Diff. คือค่าความยากของข้อสอบ

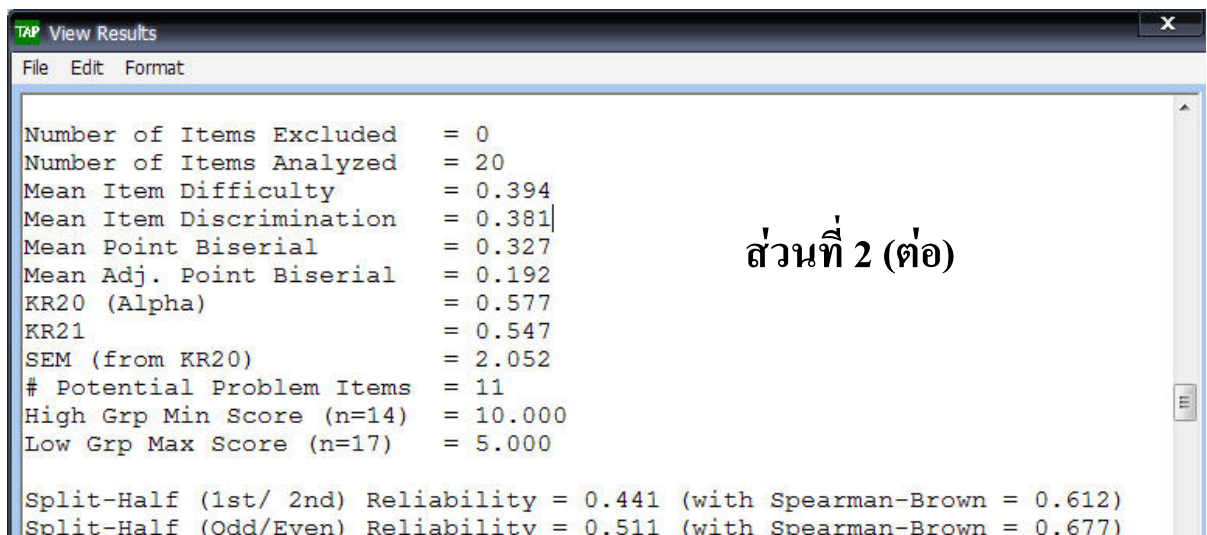
Disc. Index คือดัชนีอำนาจจำแนกของข้อสอบ

#Correct in High Grp คือจำนวนผู้สอบในกลุ่มสูงที่ตอบข้อสอบข้อนั้นถูก

#Correct in Low Grp คือจำนวนผู้สอบในกลุ่มต่ำที่ตอบข้อสอบข้อนั้นถูก

Point Biserial คือค่าอำนาจจำแนกที่คำนวณจากสูตรสหสัมพันธ์พอยท์ไบเซเรียล

Adjusted Pt. Bis. คือค่าอำนาจจำแนกที่คำนวณจากสูตรสหสัมพันธ์ไบเซเรียล



Number of Items Excluded	= 0
Number of Items Analyzed	= 20
Mean Item Difficulty	= 0.394
Mean Item Discrimination	= 0.381
Mean Point Biserial	= 0.327
Mean Adj. Point Biserial	= 0.192
KR20 (Alpha)	= 0.577
KR21	= 0.547
SEM (from KR20)	= 2.052
# Potential Problem Items	= 11
High Grp Min Score (n=14)	= 10.000
Low Grp Max Score (n=17)	= 5.000
Split-Half (1st/ 2nd) Reliability	= 0.441 (with Spearman-Brown = 0.612)
Split-Half (Odd/Even) Reliability	= 0.511 (with Spearman-Brown = 0.677)

ภาพประกอบ 13

Number of Items คือจำนวนข้อสอบทั้งหมด

Mean Item Difficulty คือค่าเฉลี่ยของค่าความยาก

Mean Item Discrimination คือค่าเฉลี่ยของดัชนีอำนาจจำแนก

Mean Point Biserial คือค่าเฉลี่ยของสหสัมพันธ์พอยท์ไบซีเรียล

KR20 (Alpha) คือค่าความเชื่อมั่นที่คำนวณด้วยสูตร KR-20

KR21 คือค่าความเชื่อมั่นที่คำนวณด้วยสูตร KR-21

SEM (from KR20) คือค่าความคลาดเคลื่อนมาตรฐานของการวัด

High Grp Min Score (n=18) คือคะแนนต่ำสุดของกลุ่มสูง

LOW Grp Max Score (n=19) คือคะแนนสูงสุดของกลุ่มต่ำ

Minimum Item Diff. คือค่าต่ำสุดของค่าความยากของข้อสอบ

Maximum Item Diff. คือค่าสูงสุดของค่าความยากของข้อสอบ

Minimum Disc. Index คือค่าต่ำสุดของดัชนีอำนาจจำแนกของข้อสอบ

Maximum Disc. Index คือค่าสูงสุดของดัชนีอำนาจจำแนกของข้อสอบ

Minimum Pt. Biserial คือค่าต่ำสุดของของสหสัมพันธ์พอยท์ไบซีเรียล

Miximum Pt. Biserial คือค่าสูงสุดของของสหสัมพันธ์พอยท์ไบซีเรียล

Item	Scale Mean if Item Deleted	Scale SD if Item Deleted	KR20 if Item Deleted	SEM if Item Deleted	Biserial Correl.	Adjusted Biserial Correl.
Item 01	7.580	3.060	0.572	2.001	0.361	0.175
Item 02	7.620	2.959	0.537	2.014	0.676	0.520
Item 03	7.620	3.143	0.595+	1.999	0.128	-0.060
Item 04	7.360	3.025	0.566	1.994	0.416	0.226
Item 05	7.640	3.071	0.573	2.008	0.356	0.174
Item 06	7.480	2.955	0.541	2.002	0.598	0.428
Item 07	7.240	2.990	0.552	2.002	0.523	0.345

ภาพประกอบ 14

Additional Item Analysis คือผลการวิเคราะห์ข้อสอบเพิ่มเติม

Item คือ ข้อสอบแต่ละข้อ

Scale Mean if Item Deleted คือค่าเฉลี่ยของข้อสอบทั้งฉบับเมื่อไม่รวมข้อสอบข้อนั้น

Scale SD if Item Deleted คือคะแนนเบี่ยงเบนมาตรฐานของข้อสอบทั้งฉบับเมื่อไม่รวมข้อสอบข้อนั้น

KR-20 if Item Deleted คือค่าความเชื่อมั่นสูตร KR-20 ของแบบทดสอบทั้งฉบับเมื่อไม่รวมข้อสอบข้อนั้น

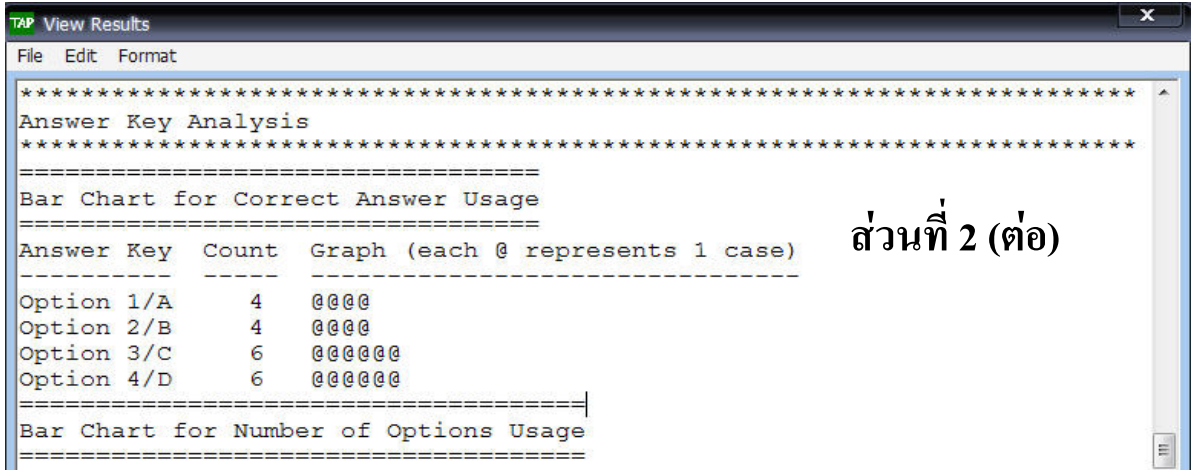
SEM if Item Deleted คือความคลาดเคลื่อนมาตรฐานของการวัดของแบบทดสอบทั้งฉบับเมื่อไม่รวมข้อสอบข้อนั้น

Biserial Correl. คือค่าสหสัมพันธ์ไบซีเรียล

Mean Biserial Corr. คือค่าเฉลี่ยของสหสัมพันธ์ไบซีเรียล

Minimum Biserial Corr. คือค่าต่ำสุดของสหสัมพันธ์ไบซีเรียล

Maximum Biserial Corr. คือค่าสูงสุดของสหสัมพันธ์ไบซีเรียล

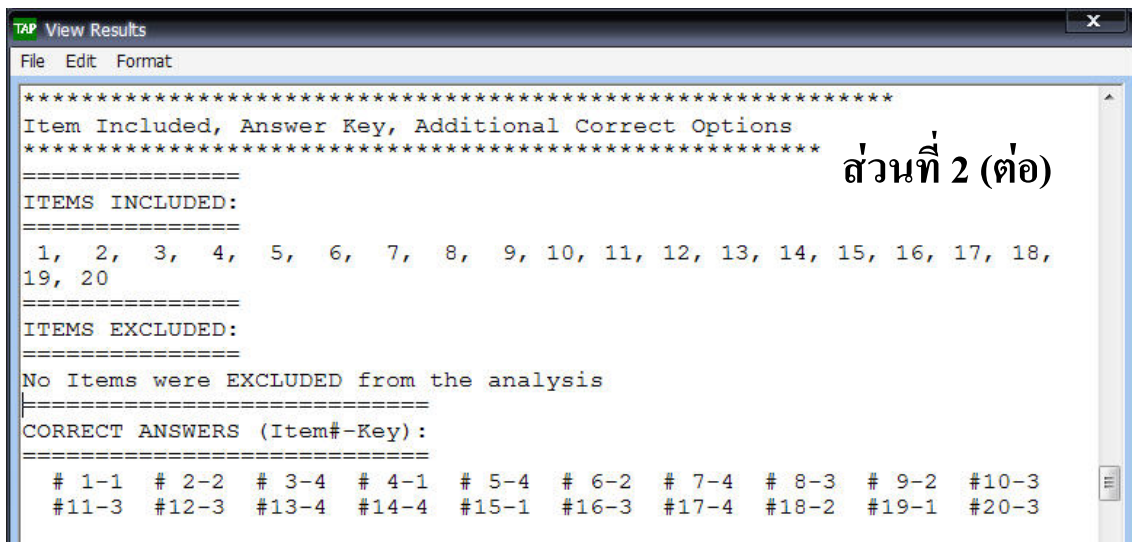


ภาพประกอบ 15

Answer Key Analysis คือการวิเคราะห์ตัวเลือกทั้งหมด

Bar Chart for Correct Answer Usage คือแผนภาพแสดงความถี่ของตัวเลือกที่เป็น
เฉลยข้อสอบ

Bar Chart for Number of Options Usage คือแผนภาพแสดงความถี่ของตัวเลือกที่ใช้
ในข้อสอบแต่ละข้อ



ภาพประกอบ 16

Item Included, Answer Key, Additional Correct Options คือแสดงข้อสอบที่รวมเข้า
วิเคราะห์ (ITEMS INCLUDED) ข้อสอบที่ไม่รวมเข้าวิเคราะห์ (ITEMS EXCLUDED:) เฉลย
ข้อสอบ และเฉลยข้อสอบ (CORRECTANSWERS (Item#-Key))

ส่วนที่ 3 ผลการวิเคราะห์รายตัวเลือก (Options Analysis) ให้ผลดังนี้

TAP View Results

File Edit Format

TITLE: MAHT-P5

Options Analysis

~~~~~ ส่วนที่ 3 ~~~~~

Item Frequencies and Percentages

page1

~~~~~

* is keyed answer, # is option that discriminates better than keyed answer

Item	Group	Option 1	Option 2	Option 3	Option 4
1	TOTAL	15* (0.300)	5 (0.100)	10 (0.200)	20 (0.400)
	High	7 (0.500)	0 (0.000)	1 (0.071)	6 (0.429)
	Low	4 (0.235)	2 (0.118)	7 (0.412)	4 (0.235)
	Diff	3 (0.265)	-2 (-0.118)	-6 (-0.340)	2 (0.193)

ภาพประกอบ 17

ตัวเลขที่อยู่นอกวงเล็บคือความถี่ของผู้สอบที่เลือกตอบตัวเลือกนั้นในแต่ละกลุ่ม ส่วนตัวเลขในวงเล็บคือสัดส่วนของความถี่นั้นต่อจำนวนผู้สอบในกลุ่ม นั้นคือในแถว Total ของตัวเลือกที่ 1 (Option 1) มีผู้สอบทั้งหมดเลือกตัวเลือก 1 จำนวน 15 คน คิดเป็นสัดส่วน 0.300 ซึ่งสัดส่วนนี้ก็คือค่าความยาก (ค่า $p = 0.30$) ของตัวเลือกที่ 1 ส่วนผลต่างของกลุ่มสูงและกลุ่มต่ำ (Diff) คือ 3 คิดเป็นสัดส่วนในวงเล็บ 0.265 ซึ่งสัดส่วนนี้ก็คือดัชนีอำนาจจำแนก (ค่า $r/D = 0.256$) ของตัวเลือกที่ 1 สังเกตว่าโปรแกรมจะคำนวณ Diff โดยใช้ความถี่ของกลุ่มสูงลบด้วยกลุ่มต่ำ ดังนั้นตัวเลขที่จำแนกได้อย่างมีประสิทธิภาพ นั้นคือกลุ่มสูงควรเลือกตอบตัวเลขน้อยกว่ากลุ่มต่ำนั่นเอง ส่วนข้อ 1 นี้ตัวเลขคือตัวเลือกที่ 1 ซึ่งมีเครื่องหมายดอกจันอยู่ มีค่าความยาก (p) 0.30 อำนาจจำแนก (r/D) 0.265 ซึ่งถือว่าเป็นข้อสอบที่ค่อนข้างยาก และมีค่าอำนาจจำแนกที่ใช้ได้ จึงเป็นข้อสอบที่ใช้ได้ ถ้าเครื่องหมาย # ก็คือตัวเลือกนั้นมีอำนาจจำแนกสูงกว่าตัวเลข

ส่วนที่ 4 การแสดงด้วยแผนภาพ (Graph : Bar Chart and Histogram) โดยด้านล่างของหน้าต่างแสดงผลลัพธ์โดยละเอียดจะปรากฏ 3 ปุ่มคือ Bar Chart for Grades แสดงแผนภูมิแท่งแจกแจงความถี่แต่ละระดับเกรด Bar Chart for Scores แสดงแผนภูมิแท่งแจกแจงความถี่แต่ละระดับคะแนน และ Histogram for Scores แสดงฮิสโตแกรมแจกแจงความถี่ระดับคะแนน ดังภาพประกอบ 19

20 TOTAL 11 (0.220) 5 (0.100) 20* (0.400) 14 (0.280)

High 1 (0.071) 3 (0.214) 6 (0.429) 4 (0.286)

Low 5 (0.294) 2 (0.118) 7 (0.412) 3 (0.176)

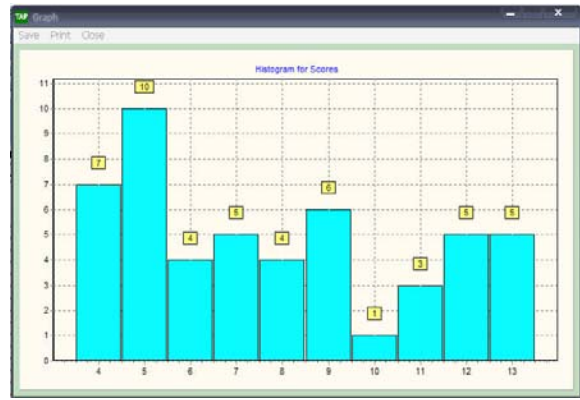
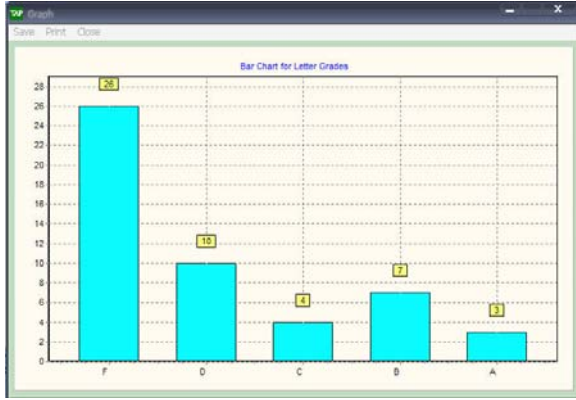
Diff -4 (-0.223) 1# (0.097) -1 (0.017) 1# (0.109)

~~~~~ ส่วนที่ 4 ~~~~~

Note that you can print this screen from the File Menu, but the formatting will be better (e.g., page breaks) if you save the file and then print from a Text Editor (such as NotePad). Any editing changes you make here WILL print, but will NOT be saved.

Bar Chart for Grades Bar Chart for Scores Histogram for Scores Close

ภาพประกอบ 18



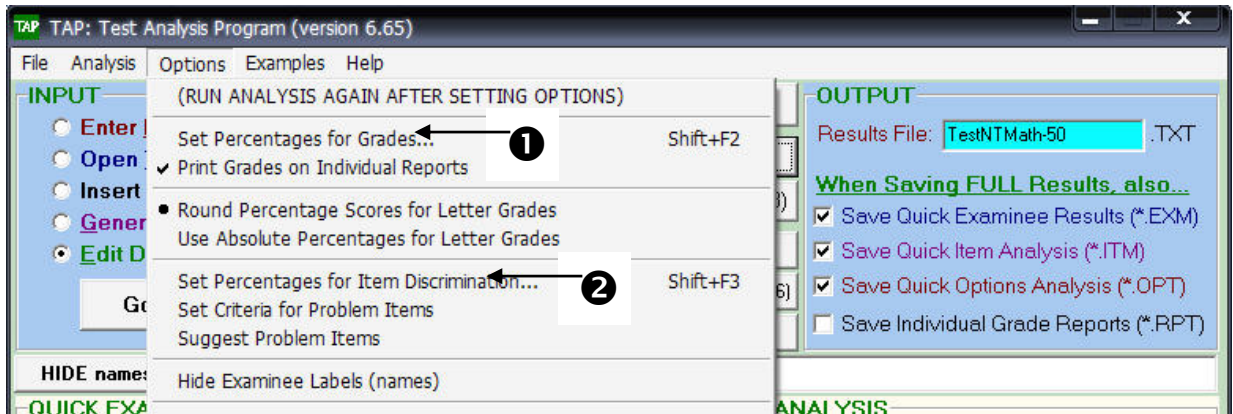
ภาพประกอบ 19

## บทที่ 5

### การเลือกใช้ เมนู Option

เมนู Options ผู้วิเคราะห์สามารถกำหนดค่าต่างในการวิเคราะห์ได้ ที่สำคัญดังนี้

1. การกำหนดค่าเปอร์เซ็นต์ของเกรด (Set Percentages If Grades )
2. การกำหนดค่าเปอร์เซ็นต์แบ่งกลุ่มสูงกลุ่มต่ำ (Set Percentages of Item Discrimination)



ภาพประกอบ 20

#### 1. การกำหนดค่าเปอร์เซ็นต์ของเกรด (Set Percentages If Grades )

Set Percentages If Grades ผู้วิเคราะห์สามารถเลือกให้ผลเกรดแก่ผู้สอบได้ โดยเลือกได้ 2 เกรด คือ ผ่านกับไม่ผ่าน เลือกได้ 5 เกรดคือ A, B, C, D และ F เลือกได้ 12 เกรดคือ A, A-, B+,B, B-, C+,C, C-, D+,D, D- และ F และไม่แสดงผลเกรด โดยแต่ละระดับเกรดสามารถเลือกเปอร์เซ็นต์ต่ำสุดที่ผู้สอบควรได้ ดังภาพประกอบ 21

- 1.1 คลิก Option ที่เมนูบาร์
- 1.2 เลือกคลิก Set Percentages If Grades ...
- 1.3 เลือกกำหนดค่า

**Grade Percentages**

**Choose the number of Grading Categories:**

2 Categories (Pass, Fail)  
 5 Categories (A, B, C, D, F)  
 12 Categories (A, A-, B+, B, B-, C+, C, C-, D+, D, D-, F)  
 User Defined:  
 No Grading Categories

**Set the minimum percentages for each Grading Category:**

| Grade | % |
|-------|---|
| A =   | 0 |
| B =   | 0 |
| C =   | 0 |
| D =   | 0 |
| F     |   |

You can change the percentages for any Grading Scheme, but you can only change the labels for the User Defined scheme (note that only the first 4 characters of your User-defined Grading Category names will print in the output).

ภาพประกอบ 21

### การเลือกกำหนดค่า

คลิกเลือก 2 Categories (Pass, Fail) = เกรด 2 ระดับ ผ่าน และไม่ผ่าน กำหนด ตัวเลขในช่อง Grade % หรือเลือกค่า Defaults ของโปรแกรม (70%)

คลิกเลือก 5 Categories (A,B,C,D,F) = เกรด 5 ระดับ A,B,C,D และF กำหนด ตัวเลข ในช่อง Grade % หรือเลือกค่า Defaults ของโปรแกรม (90,80,70,60 ตามลำดับ)

คลิกเลือก 12 Categories (A,A-,B+,B,B-,C+,C,C-,D+,D,D-,F) = เกรด 12 กำหนด ตัวเลขในช่อง Grade % หรือเลือกค่า Defaults ของโปรแกรม (94,91,88,84,81,78,74,71,68,64,61 ตามลำดับ)

คลิกเลือก User Defined = ผู้ใช้กำหนดระดับเกรดได้เอง

คลิกเลือก No Grading Categories = ไม่กำหนดเกรด

เมื่อเลือกกำหนดค่าแล้ว คลิก OK เข้าสู่หน้าจอ TAP : Test Analysis Program (Version 6.65)

ปุ่ม Analyze(F9) จะเปลี่ยนเป็น **Analyze with NEW Options (F9)** แล้วคลิก เพื่อวิเคราะห์ผลใหม่  
 สดมภ์ Ltr Grade จะแสดงผลเกรดที่กำหนด ดังภาพประกอบ 22

| HIDE names             |             | Title: MAHT-P5 (LAS-2550) |                       |                     |     |                |                                         |
|------------------------|-------------|---------------------------|-----------------------|---------------------|-----|----------------|-----------------------------------------|
| QUICK EXAMINEE RESULTS |             |                           |                       | QUICK ITEM ANALYSIS |     |                |                                         |
| ID                     | Total Score | Ltr Percent Grade         | ~68% C.I. (Raw Score) | Item                | Key | Number Correct | Item Diff. Disc. Index in High Grp in L |
| Student001             | 9           | 45.00% D                  | ( 6.9- 11.1)          | Item 01 ( 1 )#      |     | 15             | 0.30 0.26 7 (0.50) 4                    |
| Student002             | 4           | 20.00% F                  | ( 1.9- 6.1)           | Item 02 ( 2 )       |     | 13             | 0.26 0.57 8 (0.57) 0                    |
| Student003             | 4           | 20.00% F                  | ( 1.9- 6.1)           | Item 03 ( 4 )#      |     | 13             | 0.26 0.11 4 (0.29) 3                    |
| Student004             | 4           | 20.00% F                  | ( 1.9- 6.1)           | Item 04 ( 1 )#      |     | 26             | 0.52 0.48 10 (0.71) 4                   |
| Student005             | 7           | 35.00% F                  | ( 4.9- 9.1)           | Item 05 ( 4 )#      |     | 12             | 0.24 0.30 5 (0.36) 1                    |
| Student006             | 11          | 55.00% C                  | ( 8.9- 13.1)          | Item 06 ( 2 )       |     | 20             | 0.40 0.61 11 (0.79) 3                   |
| Student007             | 8           | 40.00% D                  | ( 5.9- 10.1)          | Item 07 ( 4 )       |     | 32             | 0.64 0.45 12 (0.86) 7                   |
| Student008             | 9           | 45.00% D                  | ( 6.9- 11.1)          | Item 08 ( 3 )       |     | 22             | 0.44 0.54 10 (0.71) 3                   |

ภาพประกอบ 22

## 2. การกำหนดค่าเปอร์เซ็นต์แบ่งกลุ่มสูงกลุ่มต่ำ(Set Percentages of Item Discrimination)

Set Percentages of Item Discrimination ผู้วิเคราะห์สามารถเลือกแบ่งกลุ่มสูงกลุ่มต่ำกลุ่มละกี่เปอร์เซ็นต์ก็ได้โดยระบุค่าเปอร์เซ็นต์ที่ต้องการ ดังภาพประกอบ 23

2.1 คลิก Option ที่เมนูบาร์

2.2 เลือกคลิก Set Percentages of Item Discrimination

ภาพประกอบ 23

การกำหนดจำนวนเปอร์เซ็นต์เพื่อแบ่งกลุ่มสูงกลุ่มต่ำ ในช่อง Top หรือ Bottom เป็น 33 หรือ 50 ตามความต้องการได้ หรือเลือกค่า Defaults ของโปรแกรม (27%)

เมื่อเลือกกำหนดค่าแล้ว คลิก OK เข้าสู่หน้าจอ TAP : Test Analysis Program (Version 6.65)

ปุ่ม Analyze(F9) จะเปลี่ยนเป็น **Analyze with NEW Options (F9)** แล้วคลิก เพื่อวิเคราะห์ผลใหม่

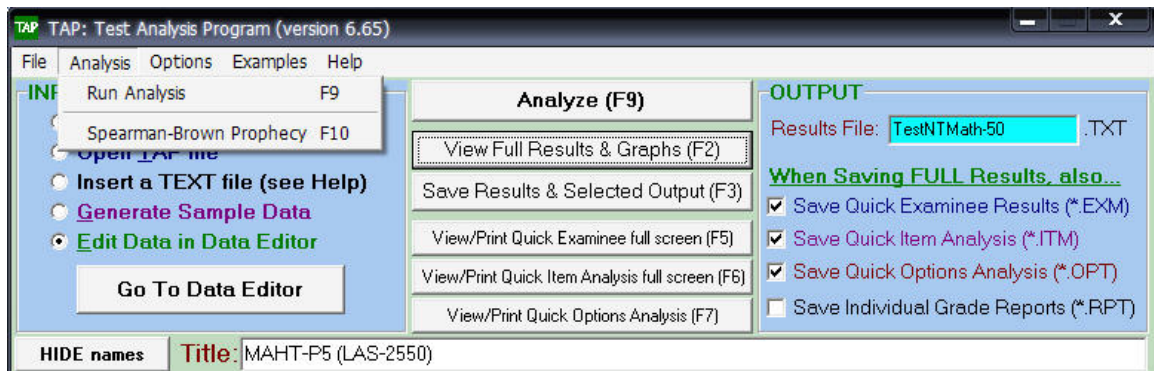
## บทที่ 6

### การเลือกใช้ เมนู Analysis

#### การเลือกใช้ เมนู Analysis

การกำหนดจำนวนข้อสอบเพิ่มเติมตามค่าความเชื่อมั่นที่ต้องการ

เมนู Analysis ประกอบไปด้วยเมนูรอง Run Analysis ซึ่งสามารถคลิกเลือกได้จากปุ่มอยู่แล้ว แต่เมนูรองที่สำคัญก็คือ Spearman-Brown Prophecy เป็นการทำนายจำนวนข้อที่ควรเพิ่มเข้าไปในแบบทดสอบเพื่อให้ได้ค่าความเชื่อมั่นตามที่กำหนด ดังภาพประกอบ 23



ภาพประกอบ 23

1. คลิก Analysis ที่เมนูบาร์
2. เลือกคลิก Spearman-Brown Prophecy แสดงหน้าจอ Spearman-Brown Prophecy ดังภาพประกอบ 24
3. แสดงคุณภาพแบบทดสอบฉบับนี้มีข้อสอบอยู่ 20 ข้อ มีความเชื่อมั่นอยู่ 0.577

**Spearman-Brown Prophecy Formula**

The reliability of the test analyzed was .....

To achieve a desired of reliability of .....

you need .....

items. This is because the test must be ...

times longer than its current .....

items, for a **TOTAL** of .....

items of similar quality to those in the test now.

These estimates are based on the formula:

**Number of Items Needed = L \* (# Items Currently in Test)**

where  $L = \frac{\text{Desired KR20} * (1 - \text{KR20})}{\text{KR20} * (1 - \text{Desired KR20})}$

ภาพประกอบ 24

จากภาพประกอบ 24 แสดงคุณภาพแบบทดสอบฉบับนี้มีข้อสอบอยู่ 20 ข้อ มีความเชื่อมั่นอยู่เดิม 0.577 ถ้าต้องการแบบทดสอบที่มีความเชื่อมั่นเพิ่มขึ้นควรจะมีข้อสอบอีกจำนวนเท่าใด

1. กำหนดตัวเลขค่าความเชื่อมั่นเพิ่มขึ้นใน To achieve a desired of reliability of ...
2. คลิกการคำนวณ Calculate

แสดงหน้าจอจำนวนข้อสอบเพิ่มเติม ดังภาพประกอบ 25

**Spearman-Brown Prophecy Formula**

The reliability of the test analyzed was ....

To achieve a desired of reliability of .....

you need .....  **more**

items. This is because the test must be ...

times longer than its current .....

items, for a **TOTAL** of .....

items of similar quality to those in the test now.

These estimates are based on the formula:

**Number of Items Needed = L \* (# Items Currently in Test)**

where  $L = \frac{\text{Desired KR20} * (1 - \text{KR20})}{\text{KR20} * (1 - \text{Desired KR20})}$

ภาพประกอบ 25

### ผลการวิเคราะห์

ค่าความเชื่อมั่นของแบบทดสอบเดิมทั้งฉบับ ( The reliability of the test analyzed was....)= 0.577

ต้องการผลค่าความเชื่อมั่นเพิ่มเติม (To Achieve a desired of reliability of ...) = 0.800

จำนวนข้อสอบที่ต้องเพิ่ม (you need....) = 39 ข้อ

ข้อสอบฉบับนี้เพิ่มมากขึ้น ( item. This is because the test must be....) = 2.94 เท่าของข้อสอบเดิม

จำนวนข้อสอบรวมทั้งหมด ( item. For a TOTAL of ...) 59 ข้อ

นั่นคือแบบทดสอบฉบับนี้มีข้อสอบอยู่เดิม 20 ข้อ มีความเชื่อมั่นอยู่เดิม 0.577 ถ้าต้องการแบบทดสอบที่มีความเชื่อมั่น 0.800 ควรจะต้องเพิ่มข้อสอบอีกจำนวนเท่าใด ผลการวิเคราะห์ก็คือ ต้องเพิ่มอีก 39 ข้อหรือ 2.94 เท่าของข้อสอบเดิม จึงจะได้ค่าความเชื่อมั่น 0.800 ข้อสอบในแบบทดสอบควรมีจำนวน 59 ข้อ ซึ่งเป็นผลการทำนายถ้าข้อสอบที่เพิ่มเติมมีคุณภาพเท่ากับข้อสอบเดิมหรือใกล้เคียงกัน



## บทที่ 7

### การตรวจสอบคุณภาพเครื่องมือวัดผลการศึกษา

เครื่องมือวัดที่นำไปใช้ในการวัดผลการเรียนรู้ของผู้เรียน ควรเป็นเครื่องมือที่ดีมีคุณภาพ เพื่อให้สามารถวัดผลการเรียนรู้ที่แท้จริงของผู้เรียนได้ เครื่องมือวัดผลที่มีคุณภาพ มีลักษณะสำคัญดังนี้

1. ความตรงหรือความเที่ยงตรง (Validity)
2. ความยากหรือความง่าย (Difficulty)
3. อำนาจจำแนก (Discrimination Power)
4. ความเที่ยงหรือความเชื่อมั่น (Reliability)
- 5.ความเป็นปรนัย (Objectivity)
6. ความสามารถในการนำไปใช้ (Usability)

**1. ความตรงหรือความเที่ยงตรง (Validity)** ความหมายของความตรงแบ่งได้เป็น 2 แนวคิด ตามแนวคิดเดิม ความตรง หมายถึง ความสามารถของเครื่องมือที่วัดได้ตรงกับสิ่งที่ต้องการวัด ส่วนแนวคิดปัจจุบัน ความตรง หมายถึง ความสามารถของหลักฐานที่อธิบายผลการวัด ได้ตรงกับสิ่งที่ต้องการวัด ความตรงของผลการวัด อาจได้มาจาก การรวบรวมหลักฐานจากหลายวิธี การตรวจสอบหลักฐานที่แสดงถึงความตรงจึงมีหลายชนิด ซึ่งจำแนกตามหลักฐานได้ 3 ชนิด คือ หลักฐานความตรงเชิงเนื้อหา หลักฐานความตรงเชิงเกณฑ์ และหลักฐานความตรงเชิงโครงสร้าง

**2. ความยากหรือความง่าย (Difficulty)** หมายถึงความยากในเชิงเนื้อหาของเครื่องมือประเภทแบบทดสอบ ซึ่งมีคำตอบถูกผิดตามหลักวิชา หรือหลักการอื่นๆ ความยากพิจารณาจากสัดส่วนหรือร้อยละของผู้ที่ตอบคำถามข้อนั้นๆ ไม่ถูกต้อง เครื่องมือวัดผลที่ดี ควรมีความยากพอเหมาะกับกลุ่มผู้สอบ สำหรับเครื่องมือประเภทที่ไม่มีคำตอบถูกผิด จะไม่มีความยากในเชิงเนื้อหา การสร้างเครื่องมือต้องใช้ภาษาที่เหมาะสมกับกลุ่มสอบ

**3. อำนาจจำแนก (Discrimination Power)** หมายถึง ความสามารถของเครื่องมือที่วัดได้ตรงกับความสามารถของผู้ถูกวัด เป็นความสัมพันธ์ระหว่างความสามารถของผู้ตอบกับการตอบถูก กล่าวคือ คนที่มีความสามารถในสิ่งที่วัดมากควรตอบถูกมาก และในทำนองเดียวกันคนที่มีความสามารถต่ำ จะตอบถูกน้อย เครื่องมือวัดที่มีอำนาจจำแนกสามารถจำแนกผู้ที่มีความสามารถแตกต่างกัน ออกจากกันตามระดับความสามารถ เช่น จำแนกคนที่รอบรู้ ออกจากที่ไม่รอบรู้ จำแนกคนที่มีคุณธรรมสูง ออกจากคนที่มีคุณธรรมต่ำ เป็นต้น

**4. ความเที่ยงหรือความเชื่อมั่น (Reliability)** หมายถึง ความคงที่ หรือความคงเส้นคงวา (Consistency) ของผลที่ได้จากการวัดสิ่งเดียวกันด้วยเครื่องมือเดียวกันในช่วงเวลาที่ต่างกัน

2 ครั้งขึ้นไป หรือความสอดคล้องของผลการวัดสิ่งเดียวกันด้วยเครื่องมือหรือวิธีการที่แตกต่างกันมากกว่าหนึ่งอย่างขึ้นไป การตรวจสอบความเที่ยงทำได้หลายวิธี เช่น วิธีวัดซ้ำ (test-retest) วิธีใช้เครื่องมือที่สมมูลกัน (equivalent forms) วิธีหาความสอดคล้องภายใน เป็นต้น

**5. ความเป็นปรนัย (Objectivity)** หมายถึง ความชัดเจนของเครื่องมือทั้งคำถามและคำตอบ คำถามที่เป็นปรนัยต้องมีความชัดเจน ผู้อ่านเข้าใจตรงกันว่าถามอะไร ส่วนคำตอบที่เป็นปรนัยต้องมีคำตอบถูกชัดเจน มีวิธีการตรวจและเกณฑ์การให้คะแนน รวมทั้งการแปลความหมายที่ผู้ตรวจสามารถให้คะแนนได้ตรงกัน

**6. ความสามารถในการนำไปใช้ (Usability)** ลักษณะของเครื่องมือที่ควรพิจารณาอีกอย่างหนึ่ง คือ ความสามารถนำไปใช้ได้สะดวก เนื่องจากครูผู้สอนหรือผู้ใช้เครื่องมือไม่ใช่ผู้ที่ได้รับการฝึกอบรมด้านการวัดผลมาโดยเฉพาะ ดังนั้นเครื่องมือจึงควรมีลักษณะที่สามารถนำไปใช้ได้สะดวก ดังนี้

6.1 ง่ายในการดำเนินการสอบ เครื่องมือควรมีคำชี้แจงสำหรับการสอบที่ง่ายและชัดเจน

6.2 ใช้เวลาในการสอบเหมาะสม เครื่องมือที่มีประสิทธิภาพเป็นเครื่องมือที่วัดได้ตรง โดยใช้คำถามจำนวนน้อยข้อ และใช้เวลาน้อยที่สุด เครื่องมือที่วัดได้ตรงต้องครอบคลุมและเป็นตัวแทนสิ่งที่วัด การใช้คำถามน้อยข้ออาจไม่ครอบคลุมสิ่งที่จะวัด การใช้คำถามมากข้ออาจทำให้ครอบคลุมสิ่งที่ต้องการวัดแต่ก็ต้องใช้เวลาการตอบเพียงพอ การใช้เวลาน้อยเกินไป อาจทำให้ผู้ตอบทำไม่ทันซึ่งอาจตอบโดยการเดา ในทางตรงกันข้ามการให้เวลามากเกินไปอาจทำให้ผู้ตอบตอบถูกเกือบทั้งหมด ซึ่งมีผลต่อค่าอำนาจจำแนกของเครื่องมือ และส่งผลให้เครื่องมือมีความเที่ยงต่ำ

6.3 ง่ายในการตรวจให้คะแนน เครื่องมือที่ดีควรมีวิธีการตรวจให้คะแนนที่สะดวก ถูกต้องไม่ซับซ้อน และใช้เวลาไม่มากนัก

6.4 ง่ายในการแปลผลและการนำผลไปใช้ เครื่องมือที่ดีควรมีคู่มือการแปลผลและการนำผลไปใช้ให้เป็นประโยชน์ต่อผู้เกี่ยวข้อง

เครื่องมือที่มีคุณภาพมีลักษณะสำคัญ หลายประการดังกล่าวแล้ว ในการสร้างเครื่องมือควรมีการตรวจสอบเครื่องมือที่สร้างขึ้นว่ามีคุณภาพหรือไม่ เครื่องมือที่จะนำไปใช้ในการวัดและประเมินผลการศึกษาควรเป็นเครื่องมือที่มีคุณภาพ ตัวชี้วัดคุณภาพของเครื่องมือ ได้แก่

1. ตัวชี้วัดคุณภาพของเครื่องมือวัดผลเป็นรายชื่อ ประกอบด้วย

ความยาก

อำนาจจำแนก

2. ตัวชี้วัดคุณภาพของเครื่องมือวัดผลทั้งฉบับ ประกอบด้วย

ความตรง

ความเที่ยง

## 1. การตรวจสอบความตรงเครื่องมือวัด

สถิติที่ใช้หาความตรง ในที่นี้จะกล่าวถึงการหาความตรงเชิงเนื้อหา 2 วิธี คือ

1. **IOC** : Index of Item – Objective Congruence
2. **CVR** : Content Validity Ratio

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน : จะตรวจสอบความตรงเชิงเนื้อหาของแบบทดสอบ  
ความสอดคล้องของข้อสอบกับวัตถุประสงค์ที่วัด  
และความครอบคลุมของเนื้อหาที่วัด

แบบสอบถามหรือแบบประเมินต่างๆ : จะตรวจสอบความสอดคล้องและครอบคลุมของ  
คำถาม(Items) กับเนื้อหา(Content) คุณลักษณะ  
(Trait)หรือองค์ประกอบของสิ่งที่ต้องการวัด

### 1.1 การหาค่า IOC พิจารณาจากคะแนนของผู้เชี่ยวชาญ

โดยการให้ตรวจผลการพิจารณาของผู้เชี่ยวชาญต่อข้อคำถามรายชื่อ 3 ประเด็น คือ  
ให้ +1 ถ้าแน่ใจว่า ข้อคำถามหรือข้อความสอดคล้องกับวัตถุประสงค์ที่ต้องการวัด  
ให้ 0 ถ้าไม่แน่ใจว่า ข้อคำถามหรือข้อความสอดคล้องกับวัตถุประสงค์ที่ต้องการวัด  
ให้ -1 ถ้าแน่ใจว่า ข้อคำถามหรือข้อความไม่สอดคล้องกับวัตถุประสงค์ที่ต้องการวัด  
นำคะแนนของผู้เชี่ยวชาญทุกคนที่ประเมินมารวมลงในแบบวิเคราะห์ความสอดคล้องของข้อคำถาม  
กับจุดประสงค์เพื่อหาค่าเฉลี่ย สำหรับข้อคำถามแต่ละข้อใช้สูตร ดังนี้

$$IOC = \frac{\sum R}{N}$$

IOC คือ ค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์  
(Index of Item – Objective Congruence)

$\sum R$  คือ ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญ

N คือ จำนวนผู้เชี่ยวชาญ

**การแปลความหมายดัชนีความสอดคล้องใช้เกณฑ์**

ถ้า  $IOC \geq 0.5$  แสดงว่า ข้อคำถามหรือข้อความวัดได้ตรงตามเนื้อหาและ  
สอดคล้องกับวัตถุประสงค์

$IOC < 0.5$  แสดงว่า ข้อคำถามหรือข้อความวัดไม่ตรงตามเนื้อหาและ  
ไม่สอดคล้องกับวัตถุประสงค์

ตัวอย่างแบบตรวจสอบความสอดคล้องของข้อสอบกับจุดประสงค์ (รายบุคคล)  
 คำชี้แจง โปรดพิจารณาข้อสอบแต่ละข้อที่แนบมาให้ว่า วัดได้ตรงกับจุดประสงค์หรือไม่

ให้กาเครื่องหมาย ✓ ลงในช่อง +1 ถ้าวัดได้ตรงกับจุดประสงค์  
 ในช่อง 0 ถ้าไม่แน่ใจหรือตัดสินใจไม่ได้  
 ในช่อง -1 ถ้าวัดได้ไม่ตรงกับจุดประสงค์

พร้อมทั้งความคิดเห็นและข้อเสนอแนะ

| จุดประสงค์ | ข้อสอบ<br>ข้อที่ | ความคิดเห็นของผู้เชี่ยวชาญ |   |    | ความคิดเห็น/<br>ข้อเสนอแนะ |
|------------|------------------|----------------------------|---|----|----------------------------|
|            |                  | +1                         | 0 | -1 |                            |
|            |                  |                            |   |    |                            |

แบบฝึกการคำนวณหาค่าดัชนีความสอดคล้องระหว่างข้อสอบแต่ละข้อกับจุดประสงค์

(Index of Item Objective Congruence : IOC)

กิจกรรม 1.1 เมื่อนำข้อสอบ จำนวน 5 ข้อ ซึ่งวัดจุดประสงค์ข้อ 1 ไปให้ผู้เชี่ยวชาญ  
 เนื้อหาวิชาพิจารณาลงความเห็นว่าคุณข้อสอบแต่ละข้อ วัดพฤติกรรมตามที่ระบุ  
 ไว้ในจุดประสงค์ข้อ 1 หรือไม่ โดยใช้แบบตรวจสอบความสอดคล้องของ  
 ข้อสอบกับจุดประสงค์ ดังนี้

| จุดประสงค์<br>ข้อที่ | ข้อสอบ<br>ข้อที่ | คะแนนความคิดเห็นของผู้เชี่ยวชาญ |         |         |         |         | ΣR    | IOC   |
|----------------------|------------------|---------------------------------|---------|---------|---------|---------|-------|-------|
|                      |                  | คนที่ 1                         | คนที่ 2 | คนที่ 3 | คนที่ 4 | คนที่ 5 |       |       |
| 1                    | 1                | +1                              | +1      | +1      | +1      | 0       | ..... | ..... |
|                      | 2                | +1                              | +1      | 0       | +1      | 0       | ..... | ..... |
|                      | 3                | -1                              | -1      | 0       | -1      | -1      | ..... | ..... |
|                      | 4                | 0                               | 0       | 0       | 0       | 0       | ..... | ..... |
|                      | 5                | +1                              | +1      | +1      | +1      | +1      | ..... | ..... |

สรุปผลการพิจารณาตรวจสอบ .....

.....

**กิจกรรม 1.2** ผู้เชี่ยวชาญ 5 คน พิจารณาแบบประเมินทักษะการทำงาน โดยมีรายการประเมิน 6 รายการ ผลการประเมินแสดงดังตาราง จงหาค่าดัชนีความสอดคล้องของรายการแต่ละรายการ และสรุปผลการพิจารณาแบบประเมินของผู้เชี่ยวชาญ

| รายการ                                                                 | ความคิดเห็นของผู้เชี่ยวชาญ |   |     |
|------------------------------------------------------------------------|----------------------------|---|-----|
|                                                                        | + 1                        | 0 | - 1 |
| 1.ด้านการวางแผนการทำงาน                                                |                            |   |     |
| 1.1 การวิเคราะห์งาน.....                                               | 5                          | 0 | 0   |
| 1.2 กำหนดเป้าหมายในการทำงาน.....                                       | 4                          | 0 | 1   |
| 1.3 กำหนดขั้นตอน กิจกรรมทำงาน                                          | 5                          | 0 | 0   |
| 1.4 กำหนดวิธีการควบคุมผลการปฏิบัติงาน                                  | 4                          | 1 | 0   |
| 1.5 บอกแหล่งข้อมูลสารสนเทศต่างๆ ที่ใช้ประกอบการตัดสินใจในการทำงาน..... | 3                          | 1 | 1   |
| 1.6 การประมาณการใช้งบประมาณ เวลา แรงงาน และวัสดุอุปกรณ์ในการทำงาน..... | 3                          | 1 | 1   |

จากสูตร  $IOC = \frac{\sum R}{N}$

- ค่าดัชนีความสอดคล้อง ข้อ 1.1 .....
- ค่าดัชนีความสอดคล้อง ข้อ 1.2 .....
- ค่าดัชนีความสอดคล้อง ข้อ 1.3 .....
- ค่าดัชนีความสอดคล้อง ข้อ 1.4 .....
- ค่าดัชนีความสอดคล้อง ข้อ 1.5 .....
- ค่าดัชนีความสอดคล้อง ข้อ 1.6 .....

สรุปผลการพิจารณาตรวจสอบ .....

.....

.....

.....

## 1.2 การคำนวณหาค่า CVR : Content Validity Ratio

จะมีตารางกำหนดมาให้ และใช้สูตรคำนวณ นำค่าไปเปรียบเทียบกับตาราง รายละเอียดดังนี้  
ตาราง ค่าอัตราส่วนความตรงเชิงเนื้อหาต่ำสุดที่ผ่านเกณฑ์ของลอรีซี

| จำนวนผู้เชี่ยวชาญ | ค่า CVR ต่ำสุด<br>ที่มีนัยสำคัญ | จำนวนผู้เชี่ยวชาญ | ค่า CVR ต่ำสุด<br>ที่มีนัยสำคัญ |
|-------------------|---------------------------------|-------------------|---------------------------------|
| 5                 | 0.99                            | 13                | 0.54                            |
| 6                 | 0.99                            | 14                | 0.51                            |
| 7                 | 0.95                            | 15                | 0.49                            |
| 8                 | 0.75                            | 20                | 0.42                            |
| 9                 | 0.78                            | 25                | 0.37                            |
| 10                | 0.62                            | 30                | 0.33                            |
| 11                | 0.59                            | 35                | 0.31                            |
| 12                | 0.56                            | 40                | 0.29                            |

ที่มา (Lawshe, 1975, p.568)

สูตรการคำนวณ

$$CVR = \frac{N_e - N/2}{N/2}$$

เมื่อ  $N_e$  = ผู้เชี่ยวชาญที่เห็นด้วยหรือเห็นว่าเหมาะสม

$N/2$  = ครึ่งหนึ่งของผู้เชี่ยวชาญทั้งหมด

ตัวอย่าง ถ้าข้อคำถามข้อหนึ่งมีผู้เชี่ยวชาญ 14 คน เห็นด้วย 11 คน คำนวณได้ ดังนี้

$$\begin{aligned} CVR &= \frac{11 - 7}{7} \\ &= 0.57 \end{aligned}$$

นำค่าที่คำนวณได้ไปเปรียบเทียบกับค่าในตารางที่จำนวนผู้เชี่ยวชาญ 14 คือ 0.51  
ค่าที่คำนวณได้ 0.57 ซึ่งสูงกว่าจึงผ่านเกณฑ์

## 2. การการวิเคราะห์ความยาก

ความยากของข้อสอบรายข้อ เป็นดัชนีที่ชี้วัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นๆ ถูก ความยากของข้อสอบรายข้อ(p) คำนวณได้จากสูตร

$$P = \frac{\text{จำนวนผู้สอบที่ตอบข้อสอบถูก}}{\text{จำนวนผู้สอบทั้งหมด}}$$

ตัวอย่าง จากผู้เข้าสอบทั้งหมด 40 คน มีผู้ตอบข้อสอบข้อที่ 1 ถูก 30 คน จงคำนวณค่าความยากของข้อสอบข้อที่ 1

$$\begin{aligned} P &= \frac{30}{40} \\ &= .75 \end{aligned}$$

ความยากมีค่าอยู่ระหว่าง 0 ถึง 1

ข้อสอบข้อที่ไม่มีผู้ใดตอบถูกเลย จะมีค่าความยาก เท่ากับ 0

ข้อสอบข้อที่มีผู้ตอบถูกครึ่งหนึ่งของผู้ตอบทั้งหมด จะมีค่าความยาก เท่ากับ .50 และ

ข้อสอบข้อที่มีผู้ตอบถูกทั้งหมดจะมีค่าความยาก เท่ากับ 1

การแปลความหมายค่าความยาก ให้เกณฑ์การพิจารณา ดังนี้

|             |                    |
|-------------|--------------------|
| 0.81 – 1.00 | ข้อสอบง่ายเกินไป   |
| 0.61 – 0.80 | ข้อสอบค่อนข้างง่าย |
| 0.41 – 0.60 | ข้อสอบยากปานกลาง   |
| 0.21 – 0.40 | ข้อสอบค่อนข้างยาก  |
| 0. – 0.21   | ข้อสอบยากเกินไป    |

กิจกรรม 2.1 จากตัวอย่าง จงคำนวณค่าความยากของข้อสอบและแปลความหมายค่าความยากข้อที่ 1-5

| ข้อสอบข้อที่ | จำนวนผู้ตอบถูก | จำนวนผู้ตอบทั้งหมด | ความยาก | ความหมาย |
|--------------|----------------|--------------------|---------|----------|
| 1            | 7              | 10                 |         |          |
| 2            | 30             | 30                 |         |          |
| 3            | 20             | 40                 |         |          |
| 4            | 0              | 28                 |         |          |
| 5            | 10             | 50                 |         |          |

สรุปคุณภาพข้อสอบ

.....  
.....

ในกรณีที่มีผู้เข้าสอบจำนวนมาก(มากกว่า 40 คน) การวิเคราะห์ความยากของข้อสอบรายข้อ อาจใช้วิธีแบ่งกลุ่มผู้สอบตามคะแนนรวม แล้วนำผลการตอบของผู้สอบที่ได้คะแนนรวมสูง ร้อยละ 27 และกลุ่มผู้สอบที่ได้คะแนนรวมต่ำ ร้อยละ 27 (เทคนิค 27%) มาวิเคราะห์ โดยใช้สูตรการคำนวณ ค่าความยากรายข้อดังนี้

$$P = \frac{H + L}{N_H + N_L}$$

H แทน จำนวนผู้ตอบถูกในกลุ่มสูง      N<sub>H</sub> แทน จำนวนผู้ตอบทั้งหมดในกลุ่มสูง  
L แทน จำนวนผู้ตอบถูกในกลุ่มต่ำ      N<sub>L</sub> แทน จำนวนผู้ตอบทั้งหมดในกลุ่มต่ำ

**กิจกรรมที่ 2.2** จากผู้เข้าสอบทั้งหมด 100 คน แบ่งเป็นผู้สอบที่ได้คะแนนรวมสูง 27 คน และผู้สอบที่ได้คะแนนรวมต่ำ 27 คน เมื่อแจกนับจำนวนผู้ตอบข้อสอบข้อที่ 2 ปรากฏว่าผู้สอบในกลุ่มสูงตอบถูก 25 คน และผู้สอบในกลุ่มต่ำตอบข้อสอบถูก 15 คน จงคำนวณค่าความยากของข้อสอบข้อที่ 2

จากโจทย์      H = ..... L = ..... N<sub>H</sub> = ..... N<sub>L</sub> = .....

จากสูตร

$$P = \frac{H + L}{N_H + N_L}$$
$$= \frac{\quad + \quad}{\quad + \quad}$$
$$= \dots\dots\dots$$

ดังนั้น ความยากของข้อสอบข้อที่ 2 เท่ากับ .....

**การวิเคราะห์ข้อสอบแบบอัตนัย**

ข้อสอบแบบอัตนัย (Essay) เป็นข้อสอบที่ให้ผู้สอบเขียนตอบตามความคิดของตนเอง เป็นเครื่องมือที่จำเป็นในการวัดผล โดยเฉพาะในการวัดเกี่ยวกับความคิดเห็น เจตคติ ความสามารถในการคิดวิเคราะห์ (Synthesis) ความสามารถในการอธิบายให้คนอื่นเข้าใจ ความสามารถในการบูรณาการความรู้ การตอบข้อสอบชนิดนี้ได้ต้องอาศัยความรู้ ประสบการณ์ และความสามารถในการบูรณาการความรู้



**วิธีการวิเคราะห์ข้อสอบแบบอัตนัย**

1. ตรวจสอบให้คะแนนข้อสอบแต่ละข้อ แล้วรวมคะแนนทุกข้อ
2. เรียงกระดาษคำตอบจากคะแนนสูงสุดมาหาค่าต่ำสุด นำจำนวนผู้ตอบจากคะแนนสูงสุดลงมา 25% เป็นกลุ่มสูง และจำนวนผู้ตอบจากคะแนนต่ำสุด 25% เป็นกลุ่มต่ำ ในการวิเคราะห์ใช้เฉพาะผลการสอบของกลุ่มสูงและกลุ่มต่ำ ที่เหลือกลุ่มกลาง 50% ไม่นำมาวิเคราะห์
3. บันทึกคะแนนของแต่ละคนในแต่ละข้อ ลงในตาราง โดยแยกตามกลุ่ม รวมคะแนนของแต่ละข้อของแต่ละกลุ่ม
4. วิเคราะห์หาค่าความยากและอำนาจจำแนก โดยใช้สูตรดังนี้

ค่าความยาก = 
$$\frac{\text{ผลรวมของคะแนนกลุ่มสูง} + \text{ผลรวมของคะแนนกลุ่มต่ำ}}{\text{จำนวนคนทั้งสองกลุ่ม} \times \text{คะแนนเต็ม}}$$

หรือ 
$$P = \frac{\sum H + \sum L}{2NM}$$

ค่าอำนาจจำแนก = 
$$\frac{\text{ผลรวมของคะแนนกลุ่มสูง} - \text{ผลรวมของคะแนนกลุ่มต่ำ}}{\text{จำนวนคนในแต่ละกลุ่ม} \times \text{คะแนนเต็ม}}$$

หรือ 
$$D = \frac{\sum H - \sum L}{NM}$$

**3. การการวิเคราะห์ค่าอำนาจจำแนก**

**3.1 ารวิเคราะห์อำนาจจำแนกของแบบทดสอบแบบอิงกลุ่ม**

สูตร

$$D = P_H - P_L$$

- เมื่อ D แทน อำนาจจำแนกของข้อสอบรายข้อ  
 P<sub>H</sub> แทน สัดส่วนของผู้ตอบข้อสอบถูกในกลุ่มสูง  
 P<sub>L</sub> แทน สัดส่วนของผู้ตอบข้อสอบถูกในกลุ่มต่ำ

หรือ

$$D = r = \frac{H - L}{N_H \text{ หรือ } N_L}$$

- เมื่อ D แทน อำนาจจำแนกของข้อสอบรายข้อ
- H แทน จำนวนผู้ตอบถูกในกลุ่มสูง
- L แทน จำนวนผู้ตอบถูกในกลุ่มต่ำ
- $N_H$  แทน จำนวนผู้ตอบทั้งหมดในกลุ่มสูง
- $N_L$  แทน จำนวนผู้ตอบทั้งหมดในกลุ่มต่ำ

**ตัวอย่าง** จากผู้เข้าสอบทั้งหมด 100 คน แบ่งเป็น 3 กลุ่ม กลุ่มที่ได้คะแนนรวมสูง 27 คน กลุ่มผู้สอนที่ได้คะแนนกลางๆ 46 คน และกลุ่มผู้สอบที่ได้คะแนนรวมต่ำ 27 คน เมื่อแจกนับจำนวนผู้ตอบข้อสอบข้อที่ 5 ปรากฏว่า ผู้สอบในกลุ่มสูงตอบถูก 25 คน และผู้สอบในกลุ่มต่ำ ตอบข้อสอบถูก 15 คน จงคำนวณค่าอำนาจจำแนกของข้อสอบข้อที่ 5

$$\text{จาก โจทย์ } P_H = \frac{25}{27} = 0.93$$

$$P_L = \frac{15}{27} = 0.56$$

$$\begin{aligned} \text{จากสูตร } D &= P_H - P_L \\ &= 0.93 - 0.56 = 0.37 \end{aligned}$$

ดังนั้น ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 เท่ากับ 0,37

**กิจกรรมที่ 3.1** จากโจทย์ตัวอย่าง จงคำนวณค่าอำนาจจำแนก ข้อที่ 5

$$H = \dots\dots\dots, L = \dots\dots\dots, N_H = N_L = \dots\dots\dots$$

จากสูตร

$$\begin{aligned} r &= \frac{H - L}{N_H \text{ หรือ } N_L} \\ &= \dots\dots\dots \\ &= \dots\dots\dots \end{aligned}$$

สรุปผลการวิเคราะห์ที่คำนวณค่าอำนาจจำแนก

.....

### 3.2 การวิเคราะห์อำนาจจำแนกของแบบทดสอบแบบอิงเกณฑ์

สูตร

$$D = P_{\text{post}} - P_{\text{pre}}$$

- D แทน อำนาจจำแนกของข้อสอบรายข้อ
- P<sub>post</sub> แทน สัดส่วนของผู้ตอบข้อสอบหลังเรียนถูก
- P<sub>pre</sub> แทน สัดส่วนของผู้ตอบข้อสอบก่อนเรียนถูก

( จากสูตร Cox and Vargs, 1966 cited in Wiersma and Jurs, 1990: 244-245)

กิจกรรมที่ 3.2 ผู้เข้าสอบ 8 จาก 25 คน ตอบข้อสอบก่อนเรียนข้อที่ 4 ได้ถูกต้อง และเมื่อสอนเสร็จ ผู้สอนให้ผู้เรียนกลุ่มเดิมตอบแบบทดสอบฉบับเดิมอีกครั้งหนึ่ง ปรากฏว่า ผู้เรียนตอบข้อสอบข้อที่ 4 ถูกต้อง 20 คน จงหาค่าอำนาจจำแนกของข้อสอบข้อที่ 4

.....

.....

.....

.....

.....

.....

### 3.3 การแปลความหมายค่าอำนาจจำแนกของข้อสอบ

อีเบลและฟรีสไบ (Ebel and Frisbi, 1986 : 234) เสนอเกณฑ์ในการพิจารณาค่าอำนาจจำแนก ดังนี้

| อำนาจจำแนก  | ผลการประเมินข้อสอบ                                         |
|-------------|------------------------------------------------------------|
| .40 ขึ้นไป  | เป็นข้อสอบที่จำแนกได้ดีมาก                                 |
| .30 - .39   | เป็นข้อสอบที่จำแนกได้ค่อนข้างดี แต่ถ้าเป็นไปได้ควรปรับปรุง |
| .20 - .29   | เป็นข้อสอบที่จำแนกได้บ้าง แต่ควรปรับปรุงแก้ไข              |
| ต่ำกว่า .19 | เป็นข้อสอบที่จำแนกได้น้อย ควรตัดออกหรือปรับปรุงใหม่        |

4. การวิเคราะห์ความเที่ยงหรือความเชื่อมั่น (Reliability)

4.1 การหาค่าความเที่ยงด้วยวิธีของคูเดอร์- ริชาร์ดสัน (Kuder – Richardson)

เป็นวิธีที่นิยมใช้กันมากเนื่องจากทดสอบกับกลุ่มตัวอย่างเพียงครั้งเดียว แบ่งย่อยเป็น 2 สูตร KR-20 และ KR-21 โดยมีข้อตกลงว่าแบบทดสอบฉบับนั้นจะต้องวัดลักษณะเดียวกันหรือวัดองค์ประกอบร่วมกัน และมีระบบการให้คะแนนเป็น Dichotomous คือ ตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน KR-21 ใช้ในกรณีเครื่องมือหรือแบบทดสอบมีความยากเท่ากันซึ่งในทางปฏิบัติเป็นไปได้ยาก จึงต้องใช้วิธีการคำนวณจากสูตร KR – 20 ดังนี้

สูตร

$$KR - 20 = \frac{k}{k - 1} \left[ 1 - \frac{\sum pq}{S^2} \right]$$

เมื่อ

- k แทน จำนวนข้อสอบ
- S<sup>2</sup> แทน ความแปรปรวนของคะแนนรวมทั้งฉบับ
- p แทน สัดส่วนของคนทำถูกในแต่ละข้อ
- q แทน สัดส่วนของคนทำผิดในแต่ละข้อ (q = 1-p)

ตัวอย่าง ผลการสอบของนักเรียน 6 คน ข้อสอบ 6 ข้อ ดังตาราง

| นักเรียน | ข้อสอบ            |      |      |      |      |      | รวมคะแนน                   |                              |
|----------|-------------------|------|------|------|------|------|----------------------------|------------------------------|
|          | 1                 | 2    | 3    | 4    | 5    | 6    | X                          | X <sup>2</sup>               |
| 1        | 1                 | 1    | 1    | 1    | 1    | 1    | 6                          | 36                           |
| 2        | 1                 | 1    | 1    | 1    | 1    | 1    | 6                          | 36                           |
| 3        | 0                 | 1    | 1    | 1    | 1    | 1    | 5                          | 25                           |
| 4        | 0                 | 0    | 1    | 1    | 1    | 1    | 4                          | 16                           |
| 5        | 1                 | 1    | 0    | 0    | 0    | 0    | 2                          | 4                            |
| 6        | 1                 | 0    | 0    | 0    | 0    | 0    | 1                          | 1                            |
| P        | 4/6               | 4/6  | 4/6  | 4/6  | 4/6  | 4/6  |                            |                              |
| Q        | 2/6               | 2/6  | 2/6  | 2/6  | 2/6  | 2/6  | <b>ΣX = 24</b>             | <b>Σ X<sup>2</sup> = 118</b> |
| pq       | 8/36              | 8/36 | 8/36 | 8/36 | 8/36 | 8/36 |                            |                              |
|          | <b>Σpq = 1.33</b> |      |      |      |      |      | <b>S<sup>2</sup> = 4.4</b> |                              |

$$KR - 20 = \frac{\dots}{\dots} \left[ 1 - \frac{\dots}{\dots} \right]$$

$$= \dots$$

**4.2 การหาความเที่ยงด้วยวิธีของครอนบัค(Cronbach)**

การหาความเที่ยงหรือความเชื่อมั่นของแบบสอบถาม นิยมใช้วิธีของครอนบัค ใช้กับแบบสอบถามที่ให้คะแนนแบบเรียงอันดับ หรือเป็นมาตราส่วนประมาณค่า (rating scale) วิธีนี้เรียกว่า การหาค่าสัมประสิทธิ์แอลฟา ( $\alpha$  -Coefficient) ซึ่งดัดแปลงมาจากสูตร KR – 20 ดังนี้

$$\alpha = \frac{n}{n-1} \left[ 1 - \frac{\sum S_i^2}{S_t^2} \right]$$

เมื่อ  $\alpha$  = สัมประสิทธิ์ความเที่ยง (สัมประสิทธิ์แอลฟา)  
 $n$  = จำนวนคำถาม  
 $\sum S_i^2$  = ผลรวมของความแปรปรวนของคะแนนในแต่ละข้อ  
 $S_t^2$  = ความแปรปรวนของคะแนนรวมทั้งฉบับ

**ตัวอย่าง** จากแบบสอบถาม แบบประมาณค่า 5 ระดับ จำนวน 6 ข้อ ผู้ตอบ 6 คน

| นักเรียน   | ข้อสอบ              |      |      |      |      |      | รวมคะแนน        |                    |
|------------|---------------------|------|------|------|------|------|-----------------|--------------------|
|            | 1                   | 2    | 3    | 4    | 5    | 6    | X               | X <sup>2</sup>     |
| 1          | 4                   | 4    | 4    | 3    | 4    | 3    | 22              | 484                |
| 2          | 4                   | 4    | 3    | 3    | 3    | 4    | 21              | 441                |
| 3          | 3                   | 3    | 2    | 4    | 3    | 3    | 18              | 324                |
| 4          | 3                   | 3    | 4    | 2    | 2    | 2    | 16              | 256                |
| 5          | 4                   | 3    | 3    | 1    | 2    | 2    | 15              | 225                |
| 6          | 2                   | 1    | 1    | 2    | 1    | 1    | 8               | 64                 |
| $\sum_i$   | 20                  | 18   | 17   | 15   | 15   | 15   | $\sum X = 100$  | $\sum X^2 = 1,794$ |
| $\sum_i^2$ | 70                  | 60   | 55   | 43   | 43   | 43   |                 |                    |
| $S_t^2$    | .67                 | 1.19 | 1.37 | 1.10 | 1.10 | 1.10 |                 |                    |
|            | $\sum S_i^2 = 6.53$ |      |      |      |      |      | $S_t^2 = 25.47$ |                    |

$$\alpha = \frac{\dots\dots}{\dots\dots} \left[ 1 - \frac{\dots\dots}{\dots\dots} \right]$$

= .....

## การพิจารณาค่าความยาก(p) และอำนาจจำแนก(r, D) แบบทดสอบ

### ตารางเกณฑ์การพิจารณาค่าความยากง่าย(p) ของตัวเลือกถูก

| ค่า p ร้อยละ | ค่า p สัดส่วน | แปลความหมาย      | ตีความหมาย   | ผลการพิจารณา            |
|--------------|---------------|------------------|--------------|-------------------------|
| 100          | 1.00          | ตอบถูกทุกคน      | ง่ายมาก      | ไม่ควร                  |
| 80           | .80           | ตอบถูก 80%       | ค่อนข้างง่าย | ถ้ามากกว่านี้ไม่ควรใช้  |
| 50           | .50           | ตอบถูกครึ่งหนึ่ง | ยากปานกลาง   | คุณภาพดีมาก             |
| 20           | .20           | ตอบถูก 20%       | ค่อนข้างยาก  | ถ้าน้อยกว่านี้ไม่ควรใช้ |
| 00           | .00           | ไม่มีคนตอบถูก    | ยากมาก       | ไม่ควรใช้               |

### ตารางเกณฑ์การพิจารณาค่าอำนาจจำแนก (r,D) ของตัวเลือกถูก

| ค่า r (D) | แปลความหมาย                           | ตีความหมาย                            | ผลการพิจารณา            |
|-----------|---------------------------------------|---------------------------------------|-------------------------|
| 100       | กลุ่มสูงถูกหมด กลุ่มต่ำผิดหมด         | จำแนกได้สูงมาก                        | เป็นข้อสอบที่มีคุณภาพ   |
| .50       | กลุ่มสูงตอบถูกมากกว่ากลุ่มต่ำ         | จำแนกได้ค่อนข้างสูง                   | คุณภาพดี                |
| .20       | กลุ่มสูงตอบถูกมากกว่ากลุ่มต่ำเล็กน้อย | จำแนกได้พอใช้                         | คุณภาพดีมาก             |
| .00       | กลุ่มสูงและกลุ่มต่ำทำถูกเท่ากัน       | จำแนกไม่ได้                           | ถ้าน้อยกว่านี้ไม่ควรใช้ |
| -1.00     | กลุ่มต่ำทำถูกหมดกลุ่มสูงทำผิดหมด      | จำแนกทางตรงข้าม(ทางลบ)ได้อย่างสมบูรณ์ | ไม่ควรใช้               |

### ตารางเกณฑ์การพิจารณาค่า p, r (D) ตัวลง

| ค่า p      | คุณภาพ    | ผลการพิจารณา | ค่า r (D)             | คุณภาพ                 | ผลการพิจารณา |
|------------|-----------|--------------|-----------------------|------------------------|--------------|
| .00 - .04  | ใช้ไม่ได้ | ไม่มีคุณภาพ  | ค่าติดลบ<br>.00 - .04 | ใช้ไม่ได้<br>ใช้ไม่ได้ | ไม่มีคุณภาพ  |
| .05 - .09  | พอใช้     | มีคุณภาพ     | .05 - .09             | พอใช้                  | คุณภาพดี     |
| .01 - .30  | ใช้ได้    |              | .10 - .30             | ใช้ได้                 |              |
| .31 - .50  | พอใช้     |              | .31 - .50             | พอใช้                  |              |
| .51 - 1.00 | ใช้ไม่ได้ | ไม่มีคุณภาพ  | .50 - 1.00            | ใช้ไม่ได้              | ไม่มีคุณภาพ  |



การอบรมเชิงปฏิบัติการวิเคราะห์ข้อมูลคุณภาพ  
เครื่องมือวัดและแบบประเมิน  
ด้วยโปรแกรม

**SPSS & TAP for Windows**

ระหว่างวันที่ 12 – 14 พฤษภาคม 2551

ห้องประชุมสี่แคว

สำนักงานเขตพื้นที่การศึกษานครสวรรค์ เขต 1



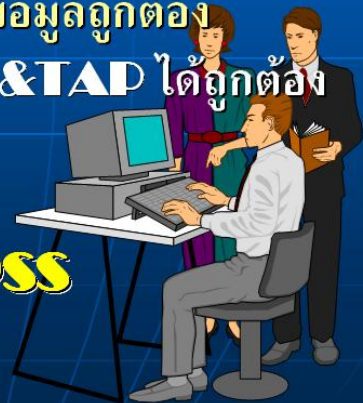
**จุดมุ่งหมายที่ต้องการ**

- \* ทบทวนการใช้สถิติเบื้องต้นถึงระดับกลาง
- \* พัฒนาความรู้ความเข้าใจการใช้สถิติวิเคราะห์คุณภาพเครื่องมือวัดและแบบประเมิน
- \* สามารถเลือกใช้สถิติวิเคราะห์ข้อมูลถูกต้อง
- \* สามารถ **Run** โปรแกรม **SPSS & TAP** ได้ถูกต้อง

อ่าน/แปลผลข้อมูล

**Print Out** จาก **SPSS**

ได้ถูกต้อง





# ต้องมีความรู้อะไรบ้างในการเรียนรู้

## SPSS & TAP for Windows

ควรมีความรู้ทางสถิติเบื้องต้นบ้าง

มีความรู้เรื่องการทดสอบ และการวิจัยบ้าง

พื้นฐานการใช้ระบบปฏิบัติการ Windows



# ที่ได้แน่ ๆ หลังการอบรม

ติดกระเป๋าคอมพิวเตอร์

ติดตัว (สมอง)

