

ความเชื่อมั่นของมาตรฐานประเมินค่าที่มีรูปแบบคำตอบ
และจำนวนมาตราต่างกัน

ปริญญานิพนธ์
ของ
ฉันทลักษณ์ ศรีผา

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและสถิติทางการศึกษา
มีนาคม 2546
ลิขสิทธิ์เป็นของ มหาวิทยาลัยศรีนครินทรวิโรฒ

371.26013

พ.254ค

จ.3

ความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบคำตอบ
และจำนวนมาตราต่างกัน

บทคัดย่อ

ของ

ฉันทลักษณ์ ศรีผา

/

24 17.3.2546

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหบัณฑิต สาขาวิชาการวิจัยและสถิติทางการศึกษา
มีนาคม 2546

ฉันทลักษณ์ ศรีผา. (2546). ความเชื่อมั่นของมาตรฐานประเมินค่าที่มีรูปแบบคำตอบและจำนวน
มาตราต่างกัน. ปรินญาณิพนธ์ กศ.ม. (การวิจัยและสถิติการศึกษา). กรุงเทพฯ :

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม :

รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์ , รองศาสตราจารย์นิภา

ศรีไพโรจน์

การวิจัยครั้งนี้ เป็นการศึกษาความเชื่อมั่นของมาตรฐานประเมินค่าที่มีรูปแบบคำตอบ
ต่างกัน 3 รูปแบบ ได้แก่ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก โดยทั้ง 3 รูปแบบ
มีจำนวนมาตรา 4 มาตรา, 5 มาตรา, 6 มาตรา, 7 มาตรา, 8 มาตรา และ 9 มาตรา
โดยมีจุดมุ่งหมายสำคัญเพื่อ 1) เพื่อเปรียบเทียบค่าความเชื่อมั่นของมาตรฐานประเมินค่า ที่มีรูปแบบ
ต่างกัน คือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จำแนกตามจำนวนมาตรา 2) เพื่อ
เปรียบเทียบค่าความเชื่อมั่นของมาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ
9 มาตรา จำแนกตามรูปแบบมาตรฐานประเมินค่า

กลุ่มตัวอย่างที่ใช้ในการศึกษา เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2
ปีการศึกษา 2545 ในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 540 คน ได้มาโดย
การสุ่มแบบหลายขั้นตอน

ผลการวิจัยพบว่า

1. การเปรียบเทียบค่าความเชื่อมั่นมาตรฐานประเมินค่าที่มีรูปแบบแตกต่างกัน คือ
แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จำแนกตามจำนวนมาตรา พบว่า ค่าความเชื่อมั่น
ของมาตรฐานประเมินค่าที่มีจำนวน 4 มาตรา และ 6 มาตรา ที่มีรูปแบบแบบกราฟฟิกมีค่าความเชื่อมั่น
สูงสุด ส่วนมาตรฐานประเมินค่าที่มีจำนวน 5, 7, 8 และ 9 มาตรา ที่มีรูปแบบแบบบรรยายมีค่า
ความเชื่อมั่นสูงสุด เมื่อใช้สูตร α และเมื่อคำนวณด้วยสูตร $r_{F.R}$ ผลที่ได้เหมือนกัน เมื่อทดสอบ
ความแตกต่างกันของค่าเชื่อมั่น โดยใช้สูตร U_{X_1} พบว่า ค่าความเชื่อมั่นมาตรฐานประเมินค่า
แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
ในเกือบทุกจำนวนมาตรา ยกเว้น เมื่อเป็นจำนวน 8 มาตรา เท่านั้นที่แตกต่างกันอย่างไม่มี
นัยสำคัญทางสถิติ เมื่อใช้สูตร α และ $r_{F.R}$ คำนวณค่าความเชื่อมั่น

2. การเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวนมาตราต่างกัน 6 มาตรา คือ จำนวนมาตรา 4, 5, 6, 7, 8 และ 9 มาตรา จำแนกตามรูปแบบมาตรประเมินค่า พบว่า ค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบแบบเดิมตัวเลขและแบบกราฟฟิกที่มีจำนวนมาตรา 8 มาตรา มีค่าความเชื่อมั่นสูงสุด เมื่อคำนวณด้วยสูตร α และ r_{F-R} คำนวณและมาตรประเมินค่าที่มีรูปแบบแบบบรรยายที่มีจำนวนมาตรา 7 และ 9 มาตรา มีค่าความเชื่อมั่นสูงสุด เมื่อคำนวณด้วยสูตร r_{F-R} และ α ตามลำดับ เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นโดยใช้สูตร UX_1 พบว่ามาตรประเมินค่า ที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เฉพาะมาตรประเมินค่าแบบบรรยายเท่านั้น เมื่อใช้สูตร α คำนวณค่าความเชื่อมั่น และพบว่า มาตรประเมินค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทุกรูปแบบ เมื่อใช้สูตร r_{F-R} คำนวณค่าความเชื่อมั่น

RELIABILITY OF RATING SCALES WITH DIFFERENT RESPONSE INTERVALS
AND FORMATS

AN ABSTRACT
BY
CHANTALUCK SRIPA

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Research and Statistics
at Srinakharinwirot University
March 2003

Chantaluck Sripa. (2003). *Reliability of Rating Scales with Different Response Intervals and Formats*. Master thesis, M.Ed. (Educational Research and Statistics)

Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee :
Assoc. Prof. Dr. Boonchird Pinyoanuntapong, Assoc. Prof. Nipa Sripairot.

This study was to investigate reliability of rating scales with different response formats, namely, filled rating scales, descriptive rating scales, and graphic rating scales and those of ratings scales were different formats of 4, 5, 6, 7, 8, and 9 intervals. The purpose of this study was two-fold. First, it aimed to compare the reliability of rating scales with different response formats as, filled rating scales, descriptive rating scales, and graphic rating scales according to their interval, and second to compare the reliability of rating scales with different response interval as, 4, 5, 6, 7, 8, and 9 intervals according to their format.

The subjects consisted of 540 Mathayom Suksa III students of secondary schools in Bangkok under the Department of General Education , during 2nd semester of 2002 academic year, through multistage random sampling.

The results revealed that :

1. The comparison the reliability of rating scales with different response formats with 4, 5, 6, 7, 8, 9 intervals resulted that those of highest reliability were ratings scales with 4 and 6 intervals and with graphic response format, whereas that of rating scale with descriptive response format was highest reliability when calculated by α and r_{F-R} formula resulting the same outcome., except that of rating scale with descriptive response format and with 4 intervals was highest reliability.

When testing the difference of the reliability by UX_1 , it was found that there was a significant different reliability at .01 level among those of rating scales with three formats and with almost all intervals, except there was no significant different among those of rating scales with three formats and with 8 intervals when calculated by α and r_{F-R} formula.

2. The comparison the reliability of rating scales with different response intervals with filled rating scales, descriptive rating scales, and graphic rating scales resulted that those of highest reliability were ratings scales with filled and graphic response format and with 8 intervals when calculated by α and r_{F-R} formula and there was a highest reliability among those of rating scales with descriptive response formats and with 7 and 9 intervals when calculated by α and r_{F-R} formula.

When testing the difference of reliability by UX_1 , it was found that there was a significant difference at .05 level among reliability of rating scales with 4, 5, 6, 7, 8 and 9 intervals and only with descriptive format when calculated by α formula, whereas there was no significant difference among those of rating scales with difference intervals when calculated by r_{F-R} formula.

ปริญญานิพนธ์
เรื่อง

ความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบคำตอบและจำนวนมาตราต่างกัน

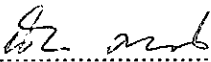
ของ
นางสาวฉันทลักษณ์ ศรีผา

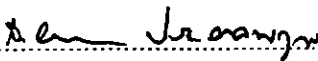
ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและสถิติทางการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

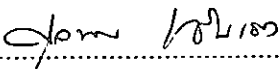

..... คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร. นภาพร หะวานนท์)
วันที่ 6 มีนาคม พ.ศ. 2546

คณะกรรมการสอบปริญญานิพนธ์


..... ประธาน
(รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์)


..... กรรมการ
(รองศาสตราจารย์ นipa ศรีไพโรจน์)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ ดร. ส.วาสนา ประवालพุก)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ดร. สุวพร เข้มเฮง)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงด้วยดี โดยได้รับความกรุณาเป็นอย่างยิ่งจาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนท์พงษ์ ที่กรุณาเสียสละเวลาอันมีค่ายิ่งในการให้คำแนะนำชี้แนะแนวทางแก้ไขข้อบกพร่อง ตลอดจนให้คำปรึกษาเกี่ยวกับการวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณอาจารย์ รศ. นิภา ศรีไพโรจน์ ที่กรุณาให้คำปรึกษา แนะนำและแก้ไขปรับปรุงข้อบกพร่อง ตลอดจนกำลังใจและความห่วงใยที่ผู้วิจัยได้รับตลอดมา

ขอกราบขอบพระคุณอาจารย์ รศ. ดร. ส.วาสนา ประवालพฤษ์ อาจารย์ ดร. สุพร เข้มแข็ง ที่เป็นกรรมการในการสอบปริญญานิพนธ์ และได้ให้คำแนะนำเพิ่มเติมเพื่อให้ปริญญานิพนธ์ฉบับนี้มีความสมบูรณ์ขึ้น

ขอกราบขอบพระคุณอาจารย์ ผศ. ระวีวรรณ พันธุ์พานิช อาจารย์ชวลิต รวยอาจิน อาจารย์วงษ์สันติ รัตนบำรุง อาจารย์เรืองอุไร คมรไชย และอาจารย์ชนม์ชกรณั์ ทินบัว ที่กรุณาให้คำปรึกษา แนะนำ ปรับปรุง ข้อบกพร่องในการตรวจเครื่องมือในการวิจัยด้วยความเอาใจใส่อย่างยิ่ง

ขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ประสิทธิ์ประสาทวิชาความรู้แก่ผู้วิจัย ตลอดจนสามารถทำให้ปริญญานิพนธ์ฉบับนี้สำเร็จด้วยดี

ขอขอบคุณผู้อำนวยการ คณะครูอาจารย์ ของโรงเรียนที่เป็นกลุ่มตัวอย่าง และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บข้อมูลเป็นอย่างดี

ขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และพี่น้องทุกคน ที่ให้การสนับสนุนและเป็นกำลังใจสำคัญยิ่งที่มีส่วนช่วยให้ปริญญานิพนธ์นี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ครั้งนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชา พระคุณบิดา-มารดา และครูอาจารย์ทุกท่าน ผู้ให้การสนับสนุนด้านการศึกษาและอบรมสั่งสอน ให้ความรู้ตลอดมา

ฉันทลักษณ์ ศรีผา

สารบัญ

บทที่		หน้า
1	บทนำ.....	1
	ภูมิหลัง.....	1
	จุดมุ่งหมายของการศึกษาค้นคว้า.....	3
	สมมติฐานของการวิจัย.....	4
	ความสำคัญของการศึกษาค้นคว้า.....	4
	ขอบเขตของการศึกษาค้นคว้า.....	4
	ประชากร.....	4
	กลุ่มตัวอย่าง.....	4
	ตัวแปรที่ใช้ในการศึกษา.....	5
	นิยามศัพท์เฉพาะ.....	5
2	เอกสารและงานวิจัยที่เกี่ยวข้อง.....	7
	มาตรฐานประเมินค่า.....	7
	รูปแบบคำตอบและจำนวนมาตราของมาตรฐานประเมินค่า.....	8
	ประโยชน์ของมาตรฐานประเมินค่า.....	20
	ความคลาดเคลื่อนจากการใช้มาตรฐานประเมินค่า.....	20
	การปรับปรุงและแก้ไขมาตรฐานประเมินค่าให้มีประสิทธิภาพ.....	22
	บุคลิภาพการทำให้ตนเองเสียเปรียบ.....	24
	ความหมายของบุคลิภาพ.....	24
	การทำให้ตนเองเสียเปรียบ.....	24
	ความเชื่อมั่น.....	30
	ความหมายของความเชื่อมั่น.....	30
	ทฤษฎีของความเชื่อมั่น.....	31
	การประมาณค่าความเชื่อมั่น.....	32
	ความคลาดเคลื่อนมาตรฐานของการวัด.....	47
	เอกสารและงานวิจัยที่เกี่ยวข้อง.....	48

สารบัญ (ต่อ)

บทที่		หน้า
3	วิธีดำเนินการศึกษาค้นคว้า.....	56
	ประชากรและกลุ่มตัวอย่าง.....	56
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	57
	การเก็บรวบรวมข้อมูล.....	65
	การวิเคราะห์ข้อมูล.....	73
	สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	74
4	ผลการวิเคราะห์ข้อมูล.....	77
	สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	77
	การเสนอผลการวิเคราะห์ข้อมูล.....	77
	ผลการวิเคราะห์ข้อมูล.....	77
5	สรุป อภิปรายผล และข้อเสนอแนะ.....	86
	จุดมุ่งหมายของการศึกษาค้นคว้า.....	86
	สมมติฐานของการศึกษาค้นคว้า.....	86
	ประชากรและกลุ่มตัวอย่าง.....	86
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	87
	การวิเคราะห์ข้อมูล.....	87
	สรุปผลการวิเคราะห์ข้อมูล.....	88
	อภิปรายผล.....	89
	ข้อเสนอแนะ.....	91
	บรรณานุกรม.....	93
	ภาคผนวก.....	99
	ประวัติย่อผู้วิจัย.....	110

บัญชีตาราง

ตาราง		หน้า
1.	แสดงจำนวนกลุ่มตัวอย่าง.....	57
2.	ค่าสถิติพื้นฐานของมาตรประเมินค่าแบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จำแนกตามจำนวนมาตรา.....	78
3.	ค่าความเชื่อมั่นของมาตรประเมินค่าทั้ง 18 ฉบับ ที่คำนวณด้วยสูตร α และ r_{F-R}	79
4.	เปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบการตอบต่างกัน จำแนกตามจำนวนมาตรา.....	81
5.	เปรียบเทียบความแตกต่างค่าความเชื่อมั่นรายคู่ของมาตรประเมินค่า ที่มีรูปแบบการตอบต่างกัน จำแนกตามจำนวนมาตรา.....	82
6.	เปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวนมาตราต่างกัน จำแนกตามรูปแบบมาตรประเมินค่า.....	84
7.	เปรียบเทียบความแตกต่างค่าความเชื่อมั่นรายคู่ของมาตรประเมินค่า แบบบรรยายที่มีจำนวนมาตราต่างกัน.....	85
8.	ค่าความสอดคล้องของมาตรประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบ...	103
9.	ค่าอำนาจจำแนกของมาตรประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบ.....	104

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1 ภาพแสดงแบบแผนการจัดกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล.....	67
--	----

บทที่ 1

บทนำ

ภูมิหลัง

การวิจัยทางการศึกษาเป็นวิธีการศึกษาหาความรู้ความจริงที่เชื่อถือได้โดยวิธีการที่มีระบบแบบแผนที่เชื่อถือได้ เพื่อนำความรู้นั้นไปสร้างกฎเกณฑ์ทฤษฎีต่าง ๆ หรือเพื่อนำไปอธิบายทำนายและควบคุมปรากฏการณ์ต่าง ๆ ซึ่งผลของการวิจัยจะมีคุณค่าเพียงใดนั้นส่วนหนึ่งขึ้นอยู่กับ การเก็บรวบรวมข้อมูลที่ติดด้วยเครื่องมือที่มีคุณภาพ นั่นคือถ้าเครื่องมือมีคุณภาพจะส่งผลให้ข้อมูล มีความเชื่อถือได้ เมื่อข้อมูลมีความเชื่อถือได้ก็จะส่งผลให้การวิจัยมีประสิทธิภาพ ดังนั้นสิ่งหนึ่งที่ ผู้วิจัยควรตระหนักอย่างยิ่งในการวิจัย คือ คุณภาพของเครื่องมือ กล่าวคือถ้าแบบประเมินใดขาด ความเที่ยงตรงและความเชื่อมั่นแล้วผลที่ได้ก็จะไม่มีความหมายแต่ประการใด (อนันต์ ศรีโสภณ. 2521 : 24) ซึ่งสอดคล้องกับคำกล่าวของ นันนอลลี (Nunnally. 1964: 59) ที่ว่าสิ่งที่จำเป็นต้อง ทราบอย่างหนึ่งของเครื่องมือ ก็คือค่าความเชื่อมั่น (Reliability) ของเครื่องมือ เพราะค่าความเชื่อมั่น เป็นดัชนีบ่งชี้ว่าเครื่องมือนั้นมีคุณภาพหรือไม่ ในการสร้างเครื่องมือทุกครั้งจึงจำเป็นต้องทำการ ตรวจหาค่าความเชื่อมั่น (บุญเชิด ภิญโญนนท์พวงษ์. 2537 : 49) ในการวัดลักษณะทางจิตวิทยามี เครื่องมือที่ใช้วัดอยู่หลายชนิดแต่ละชนิดมีคุณสมบัติแตกต่างกัน แล้วแต่การเลือกใช้ให้เหมาะสม สอดคล้องกับจุดมุ่งหมายของการวัด และเครื่องมือหนึ่งที่ได้รับการนิยมนำมาใช้อีกกัน อย่างแพร่หลาย คือมาตรประเมินค่า (Rating Scales) (Guilford. 1954 : 263) ทั้งนี้อาจเป็น เพราะมาตรประเมินค่ามีความสะดวกในการใช้ สร้างง่าย ใช้วัดพฤติกรรมได้หลายลักษณะ และ ลักษณะที่สำคัญของมาตรประเมินค่าอยู่ตรงที่คำตอบซึ่งมีลักษณะเป็นการเปรียบเทียบปริมาณ มากน้อยผู้ตอบจะต้องตอบด้วยการประเมินสถานการณ์ อันเป็นสิ่งที่เร้าที่กำหนดให้และเลือกตอบ เพียงคำตอบเดียว (บุญธรรม กิจปรีดาบริสุทธิ์. 2531 : 113) มาตรประเมินค่าจึงเป็นวิธีการ ประเมินบุคคลที่ได้รับความนิยมมากที่สุด และยังมีการนำมาใช้ในรูปของแบบสอบถามอย่างแพร่หลาย ในการวิจัย (Cunningham. 1986:338)

มาตรประเมินค่า (Rating scales) จะประกอบด้วยสองส่วน คือ ส่วนที่เป็นชุดของ ข้อความแสดงพฤติกรรมของคุณลักษณะที่ต้องการวัดกับส่วนที่เป็นชุดของคำตอบและถ้า พิจารณาจากในส่วนที่เป็นชุดของคำตอบนั้น สามารถแบ่งมาตรประเมินได้หลายรูปแบบ ดังเช่น การแบ่งตามรูปแบบคำตอบ (Response Formats) ซึ่งอาจแบ่งได้เป็น 5 รูปแบบ คือ มาตรประเมินค่าแบบตัวเลข (Numerical Scales) มาตรประเมินค่าแบบกราฟฟิก (Graphic Scales) มาตรประเมินค่าแบบมาตรฐาน (Standard Scales) มาตรประเมินค่าแบบแต้มสะสม (Cumulated Points Scales) และมาตรประเมินค่าแบบบังคับให้เลือก (Forced Choice Scales)

(Guidford. 1954 : 263-277) ซึ่งในการแบ่งตามรูปแบบคำตอบนั้นยังมิมีผู้จำแนกประเภทได้ต่างจากกิลฟอร์ดดังที่กล่าวข้างต้น เช่น อาจแบ่งเป็นรูปแบบ แบบเติมตัวเลข แบบรอยยิ้ม แบบคำตอบจำกัด แบบเปรียบเทียบรายคู่ แบบจัดอันดับ เหล่านี้เป็นต้น หรืออาจแบ่งตามลักษณะที่เป็นมาตรวัดเป็น 2 รูปแบบ คือ รูปแบบมาตรวัดขั้วเดียว (Unipolar) และมาตรวัดสองขั้ว (Bipolar) แล้วแต่ความสนใจของผู้สร้าง นอกจากนั้นในส่วนที่เป็นชุดของคำตอบของมาตรประเมินค่านั้นสามารถกำหนดจำนวนมาตราได้หลายมาตราขึ้นอยู่กับความต้องการและความสะดวกของผู้วิจัย ได้แก่ 3 มาตรา 4 มาตรา 5 มาตรา 6 มาตรา 7 มาตรา เป็นต้น

ถึงแม้ว่ามาตรประเมินค่าจะเป็นเครื่องมือที่สร้างง่ายและมีความสะดวกในการใช้กว่าเครื่องมือชนิดอื่นๆ แต่ควรใช้ด้วยความระมัดระวังเนื่องจากมีข้อจำกัดอันเกิดจากความคลาดเคลื่อนในการวัดซึ่งเป็นปัญหาสำคัญที่ทำให้มาตรประเมินค่าขาดความเชื่อถือทำให้ข้อมูลที่ได้อาจมีความเที่ยงตรงและความเชื่อมั่นต่ำ ได้มีนักจิตวิทยาหลายท่านศึกษาสาเหตุความคลาดเคลื่อนในมาตรประเมินค่าพบว่า มี 3 แหล่ง คือ ความคลาดเคลื่อนจากเครื่องมือ ความคลาดเคลื่อนจากการประเมินและความคลาดเคลื่อนจากตัวผู้ประเมิน ซึ่งจากการศึกษาพบว่าความคลาดเคลื่อนส่วนใหญ่เกิดจากผู้ประเมินและการสร้างเครื่องมือเป็นสำคัญ นักจิตวิทยาและนักการศึกษาจึงได้หาแนวทางแก้ไขเพื่อเพิ่มความเชื่อถือได้ในการใช้มาตรประเมินค่าและพบว่าแนวทางในการปรับปรุงเครื่องมือให้ดีและมีความหมายมากขึ้นโดยปรับปรุงในส่วนที่เป็นชุดของข้อความแสดงพฤติกรรมของคุณลักษณะที่ต้องการวัดกับส่วนที่เป็นชุดของคำตอบซึ่งในส่วนที่เป็นชุดของคำตอบนี้เองสามารถสร้างตามความต้องการและความสะดวกของผู้สร้าง ทำให้มีมาตรประเมินค่าหลายแบบต่างๆ กัน (Well.1960 : 393-397) ดังนั้นจึงมีการวิจัยจำนวนมากที่ต้องการตรวจสอบว่ารูปแบบของคำตอบ และจำนวนมาตราของมาตรประเมินค่ามีผลต่อการตอบมาตรประเมินค่าอย่างไรเพื่อพัฒนาและปรับปรุงมาตรประเมินค่าให้มีคุณค่าและประโยชน์สูงสุด ดังเช่น การวิจัยของพนิดา พุ่งเกียรติินาส (2540 : บทคัดย่อ) ได้ศึกษาความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบการตอบแตกต่างกัน ผลสรุปว่า มาตรประเมินค่าที่มีรูปแบบคำตอบต่างกันมีความเชื่อมั่นแตกต่างกัน ซึ่งสอดคล้องกับคำกล่าวของกิลฟอร์ดที่ว่ารูปแบบของคำตอบมีผลต่อการตอบของผู้ตอบแตกต่างกัน (Guilford. 1954 : 263) และสอดคล้องกับผลวิจัยของ นิวสเต็ด และอาร์โนล (Newstead and Arnold. 1989 : abstract) ได้ศึกษาผลของรูปแบบคำตอบของมาตรประเมินค่าที่มีต่อผลการประเมินการสอนของครู ผลปรากฏว่า มาตรประเมินค่าที่มีรูปแบบคำตอบต่างกันทั้งสามรูปแบบมีค่าเฉลี่ยแตกต่างกัน นอกจากนี้ในการกำหนดจำนวนมาตราของมาตรประเมินค่าว่าควรจะมีกี่มาตรนั้นเป็นองค์ประกอบหนึ่งที่มีต่อความเชื่อมั่นที่ควรคำนึงถึง อุทุมพร จามรมาน (2538 : 70) กล่าวว่า การกำหนดจำนวนมาตราของมาตรประเมินค่าว่าควรจะมีกี่ช่วงนั้นขึ้นอยู่กับความละเอียดของข้อมูลของผู้สร้างต้องการโดยปกติ นิยม 3 ถึง 8 มาตรา และจากการศึกษาในเรื่องจำนวนมาตราของสเกลนี้ พบว่าจำนวนช่วงที่เหมาะสมและลดความคลาดเคลื่อนได้ คือ จำนวน 3-7 มาตรา จากการศึกษาของไซมอนด์ (Symond,1925 : abstract) พบว่า

จำนวนมาตราที่เหมาะสมที่ทำมาตรประเมินค่ามีความเชื่อมั่นสูงสุด คือ 7 มาตร แต่จากการศึกษาของเบนดิง (Bending, 1953 : abstract) และกิลฟอร์ด (Guilford, 1954 : 289-299) พบว่าจำนวน 7 มาตร ไม่ได้เหมาะสมเสมอไปขึ้นอยู่กับว่าผู้ประเมินมีความรู้ในเรื่องที่ประเมินเพียงใด นอกจากนี้ ยังพบว่าจำนวนมาตรา 3, 5, 7 และ 9 มาตรามีความเชื่อมั่นไม่แตกต่างกัน และความเชื่อมั่นจะลดลงเมื่อใช้มาตรประเมินค่า 11 มาตร

จากการศึกษาค้นคว้างานวิจัยมาตรประเมินค่าพบว่าส่วนใหญ่นิยมประมาณค่าความเชื่อมั่นโดยใช้สัมประสิทธิ์แอลฟา (α) ของครอนบัค (Cronbach) คำนวณหาค่าความสอดคล้องภายในของแบบทดสอบ ในความเป็นจริงแล้วสัมประสิทธิ์แอลฟาที่คำนวณได้จะเท่ากับความเชื่อมั่นของแบบทดสอบที่ประกอบด้วยส่วนย่อยได้ก็ต่อเมื่อส่วนย่อยของแบบทดสอบเป็นไปตามข้อตกลงของคะแนนจริงสมมูล (Essentially Tau Equivalent) แต่ในทางปฏิบัติไม่อาจจะทำได้ ดังนั้นการใช้สัมประสิทธิ์แอลฟาจึงเป็นการละเมิดข้อตกลงดังกล่าว แล้วจะส่งผลทำให้สัมประสิทธิ์แอลฟามีค่าต่ำกว่าหรือสูงกว่าสัมประสิทธิ์ความเชื่อมั่นของประชากร (บุญเชิด ภิญญอนันตพงษ์. 2537 : 57-58 ; อ้างอิงจาก Zimmerman, Zumbo and Lalonde. 1993. *Educational and Psychological Measurement*) ดังนั้นในการประมาณความเชื่อมั่นของมาตรประเมินค่าควรพิจารณาให้รอบคอบ ซึ่งการประมาณค่าความเชื่อมั่นโดยใช้ข้อตกลงตามแบบจำลองคะแนนจริงสัมพันธ์จะเป็นวิธีการที่เหมาะสม เป็นวิธีที่มีความปลอดภัยหรือมีความเสี่ยงน้อยที่สุดที่จะให้ผลคลาดเคลื่อน (บุญเชิด ภิญญอนันตพงษ์. 2537 : 50 – 70)

จากที่กล่าวมาข้างต้น จะพบว่ารูปแบบมาตรประเมินค่านั้นสามารถทำได้หลายรูปแบบและเป็นเครื่องมือที่ได้รับความนิยมมากในการวัดคุณลักษณะทางจิตวิทยา แต่ยังไม่พบว่ามีมาตรประเมินค่ารูปแบบใด ที่มีจำนวนมาตราเท่าไร จึงจะมีความเชื่อมั่นมากที่สุด ทำให้ผู้วิจัยมีความสนใจที่จะศึกษาความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบคำตอบและจำนวนมาตราต่างกัน เพื่อจะศึกษาว่าในรูปแบบใดควรจะมีระดับมาตราที่มากที่สุด จึงจะมีความเชื่อมั่นมากที่สุด โดยใช้สูตร สัมประสิทธิ์แอลฟาของ ครอนบัค (Cronbach) และสูตรของ เฟลด์ ราซุ โดยเลือกมาตรประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบมาใช้ประกอบในการศึกษาค้นคว้า

จุดมุ่งหมายของการศึกษาค้นคว้า

จุดมุ่งหมายสำคัญของการศึกษาค้นคว้าครั้งนี้ เพื่อศึกษาว่ามาตรประเมินค่าที่มีรูปแบบต่างกัน 3 รูปแบบ ได้แก่ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก รูปแบบใดมีความเชื่อมั่นมากที่สุด และจะมีค่าแตกต่างกันหรือไม่ เมื่อจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มาตร

โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบต่างกัน คือ แบบเต็มตัวเลข แบบบรรยาย และแบบกราฟฟิก จำแนกตามจำนวนมาตรา
2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มาตรา จำแนกตามรูปแบบมาตรประเมินค่า

สมมติฐานของการวิจัย

1. มาตรประเมินค่าที่มีรูปแบบมาตรวัดแตกต่างกัน คือ แบบเต็มตัวเลข แบบบรรยาย และแบบกราฟฟิก มีค่าความเชื่อมั่นแตกต่างกัน
2. มาตรประเมินค่าแต่ละรูปแบบที่มีจำนวนมาตราต่างกัน 6 มาตรา คือ 4, 5, 6, 7, 8 และ 9 มาตรา มีค่าความเชื่อมั่นแตกต่างกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้จะ เป็นประโยชน์ต่อนักวิจัยนักวัดผลและผู้ที่มีส่วนเกี่ยวข้องใช้เป็นแนวทางในการตัดสินใจว่าควรเลือกใช้มาตรประเมินค่ารูปแบบใดและกี่จำนวนมาตราจึงจะเหมาะสมในการเก็บรวบรวมข้อมูลเพื่อให้ได้ข้อมูลที่มีความเชื่อมั่นสูงและการศึกษาค้นคว้าครั้งนี้ยังใช้เป็นแนวทางตัวอย่างของการศึกษาวิจัยเกี่ยวกับความเชื่อมั่น

ขอบเขตของการศึกษาค้นคว้า

1. ประชากร
ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ของโรงเรียนมัธยมศึกษา ในกรุงเทพมหานคร จำนวน 113 โรงเรียน จำนวนนักเรียน 49,836 คน
2. กลุ่มตัวอย่าง
กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ของโรงเรียนมัธยมศึกษา ในกรุงเทพมหานคร จำนวน 540 คน ซึ่งได้มาโดยวิธีสุ่มหลายขั้นตอน (Multistage Random Sampling) และประมาณกลุ่มตัวอย่างด้วยความเชื่อมั่น 95% ที่ความคลาดเคลื่อน 10% เมื่อเปรียบเทียบตารางกำหนดขนาดของกลุ่มตัวอย่างสำหรับศึกษาค่าเฉลี่ยประชากรพบว่าต้องใช้กลุ่มตัวอย่างจำนวน 398 คน (ศิริชัย กาญจนวาสี และคณะ. 2537 : 106)

3. ตัวแปรที่ใช้ในการศึกษา

3.1 ตัวแปรอิสระ ได้แก่

3.1.1 รูปแบบมาตรประเมินค่า

- แบบเติมตัวเลข
- แบบบรรยาย
- แบบกราฟฟิก

3.1.2 จำนวนมาตรา จำแนกเป็น 6 มาตรา ได้แก่ 4, 5, 6, 7,

8 และ 9 มาตรา

3.2 ตัวแปรตาม ได้แก่ ค่าความเชื่อมั่นของมาตรประเมินค่าที่คำนวณได้จากสูตรสัมประสิทธิ์แอลฟาของ ครอนบาค (Cronbach) และสูตรของ เฟลด์ท์ ราชู

นิยามศัพท์เฉพาะ

1. มาตรประเมินค่า (Rating Scales) หมายถึง เครื่องมือที่ใช้ในการวัดคุณลักษณะทางจิตวิทยาได้ค่าที่วัดเป็นค่าเชิงปริมาณอย่างเป็นระบบ โดยมีสเกลที่แบ่งเป็นจำนวนมาตราเท่าๆ กัน มาตรประเมินค่าประกอบด้วยสองส่วน คือ ส่วนที่เป็นชุดของข้อความแสดงพฤติกรรมของคุณลักษณะที่ต้องการวัดกับส่วนที่เป็นชุดของคำตอบซึ่งจะแสดงความเข้มของระดับพฤติกรรมของคุณลักษณะนั้นๆ โดยมีข้อตกลงว่าสเกลให้ค่าที่เป็นแบบจำนวนมาตราที่เท่า ๆ กัน ในงานวิจัยนี้เป็นแบบวัดลักษณะการทำให้ตนเองเสียเปรียบ

2. รูปแบบมาตรประเมินค่าของการวิจัยครั้งนี้ได้กำหนดรูปแบบ ดังนี้คือ แบบเติมตัวเลข แบบกราฟฟิก และแบบบรรยาย

2.1 มาตรประเมินค่าแบบเติมตัวเลข (Filled Rating Scale) หมายถึง มาตรประเมินค่าที่วัดลักษณะประกอบด้วยระดับพฤติกรรมทางเดียวที่เป็นบวกหรือลบอย่างใดอย่างหนึ่ง มีคำตอบเป็นตัวเลขให้ผู้ประเมินเลือกนำค่าตัวเลขที่ตรงกับระดับพฤติกรรมไปใส่หน้าข้อความที่เป็นพฤติกรรมนั้นที่ตรงกับความรู้สึกของผู้ประเมินมากที่สุด

2.2 มาตรประเมินค่าแบบบรรยาย (Descriptive Rating Scales) หมายถึง รูปแบบหนึ่งของมาตรประเมินค่า ประกอบด้วยรายการพฤติกรรม และคำตอบที่เป็นคำบรรยายหรือข้อความแสดงระดับพฤติกรรมที่ให้ผู้ประเมิน ให้ผู้ประเมินเลือกข้อความที่ตรงกับความรู้สึกของเขามากที่สุด

2.3 มาตรประเมินค่าแบบกราฟฟิก (Graphic Rating Scales) หมายถึง รูปแบบหนึ่งของมาตรประเมินค่า ประกอบด้วยรายการพฤติกรรม มีคำตอบเป็นกราฟหรือรูปภาพเพื่อชี้ค่าที่ประเมิน โดยมีข้อความกำกับหัวท้ายที่ตรงข้ามกันหรือต่างขั้วกัน

3. จำนวนมาตรา คือ การแบ่งมาตรการให้คะแนนเป็นระดับตามความเข้มต่างๆ จำแนกได้เป็น จำนวนมาตรา 6 มาตรา ได้แก่ 4, 5, 6, 7, 8 และ 9 มาตรา

4. ค่าความเชื่อมั่น หมายถึง สัดส่วนของความแปรปรวนคะแนนจริงต่อความแปรปรวนคะแนนที่สอบได้ โดยความแปรปรวนคะแนนจริงคำนวณมาจากความแปรปรวนร่วมระหว่างคะแนนส่วนย่อยของ มาตรการประเมินค่าที่สอบกับกลุ่มตัวอย่างครั้งเดียวกัน ในการศึกษาครั้งนี้ เป็นการศึกษาค่าความเชื่อมั่นที่ได้จากสูตรหาค่าความเชื่อมั่นจากสูตร r_{F-R} ของเฟลด์ - ราซ และสัมประสิทธิ์แอลฟาของครอนบัค

5. ลักษณะนิสัยการทำให้ตนเองเสียเปรียบ (Self handicapping) คือ พฤติกรรม การกล่าวอ้างให้ผู้อื่นเชื่อว่ามิบังจายที่ทำให้การกระทำของบุคคลนั้นตกอยู่ในสถานการณ์ที่จะได้รับความล้มเหลว เช่น ความเจ็บป่วย ความโชคร้าย ความเมื่อยล้า การที่ต้องทานยา การไม่พยายามอย่างเต็มที่ การรบกวนจากสิ่งอื่น การพักผ่อนไม่เพียงพอ ความไม่พร้อมและความวิตกกังวลมากเกินไป เหล่านี้เป็นต้น

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า ผู้วิจัยได้ศึกษาเนื้อหาต่างๆ โดย
เสนอรายละเอียดดังนี้

1. มาตรฐานประเมินค่า
 - ความหมายของมาตรฐานประเมินค่า
 - รูปแบบคำตอบและจำนวนมาตราของมาตรฐานประเมินค่า
 - ประโยชน์ของมาตรฐานประเมินค่า
 - ลักษณะของความคลาดเคลื่อนในมาตรฐานประเมินค่า
 - การปรับปรุงและแก้ไขการใช้มาตรฐานประเมินค่าให้มีประสิทธิภาพ
2. บุคลิกภาพการทำให้ตนเองเสียเปรียบ
 - ความหมายของบุคลิกภาพ
 - การทำให้ตนเองเสียเปรียบ
3. ความเชื่อมั่น
 - ความหมายของความเชื่อมั่น
 - ทฤษฎีความเชื่อมั่น
 - การประมาณค่าความเชื่อมั่น
 - ปัจจัยที่ส่งผลกระทบต่อความเชื่อมั่น
 - ความคลาดเคลื่อนมาตรฐานของการวัด
4. เอกสารและงานวิจัยที่เกี่ยวข้อง

1. มาตรฐานประเมินค่า (Rating scale)

ความหมายของมาตรฐานประเมินค่า

คำ "Rating Scale" นี้ นักวิชาการของไทยได้บัญญัติศัพท์เป็นภาษาไทยไว้ต่าง ๆ กัน ได้แก่ มาตราส่วนประเมินค่า แบบประเมินค่า แบบประมาณค่า มาตราวัดอันดับคุณภาพ มาตรฐานประมาณค่า และมาตรฐานประเมินค่า เป็นต้น แต่ในงานวิจัยนี้ ผู้วิจัยเลือกใช้คำ "มาตรฐานประเมินค่า" แทน "Rating Scale"

กาญจนา มณีแสง (2523 : 45) กล่าวว่า มาตรประเมินค่า เป็นเครื่องมือใช้สำหรับ ประเมินค่าของสถานการณ์หรือคุณลักษณะออกเป็นระดับต่างๆ และเป็นเครื่องมือที่ใช้วัดแนวโน้ม ของคุณลักษณะต่างๆ ในรูปของมาตราที่วัดคุณค่าเหล่านั้นออกไปเป็นจำนวนได้

วิเชียร เกตุสิงห์ (2530 :73) กล่าวว่า มาตรประเมินค่า เป็นแบบทดสอบที่มีข้อความหรือ ข้อคำถามเกี่ยวกับเรื่องให้ผู้ตอบพิจารณา พร้อมกับมีคำถามแสดงความเข้มของความคิดเห็น เกี่ยวกับ เรื่องนั้น เป็นระดับ มาก – น้อย สูง – ต่ำ ส่วนใหญ่ใช้ตั้งแต่ 3 ระดับขึ้นไป

วิรัช วรรณรัตน์ (2532 :15) กล่าวว่า มาตรประเมินค่า เป็นมาตรวัดที่ใช้ส่วนบุคคล ให้ตอบสนองในรูปของความรู้สึกหรือเจตคติที่มีต่อลักษณะ คุณค่า บุคลิกภาพ หรือความต้องการ แล้วให้นำน้ำหนักคะแนนแทนความรู้สึกหรือเจตคตินั้น

กู๊ด (Good. 1973) กล่าวว่า มาตรประเมินค่า คือ สิ่งที่ทำขึ้นเพื่อใช้ในการ ประเมิน ผลผลิต เจตคติ หรือบุคลิกลักษณะแต่ละบุคคล โดยเฉพาะลักษณะที่วัดได้ยากซึ่งผลการวัด จะได้ค่าตัวเลขที่แสดงปริมาณของลักษณะที่ถูกวัด

กิลฟอร์ด (Guilford. 1954 : 263) กล่าวว่า มาตรประเมินค่าเป็นวิธีการวัดทางจิตวิทยา (Psychological – measurement) ที่ขึ้นอยู่กับการตัดสินใจของผู้ประเมินในการประเมินพฤติกรรม บุคลิกลักษณะหรือผลผลิต

จากความหมายดังกล่าวพอสรุปได้ว่า มาตรประเมินค่า หมายถึง เครื่องมือที่ใช้ประเมิน คุณลักษณะต่างๆอันเป็นสิ่งที่ซับซ้อนและเป็นนามธรรม วัดได้ยาก ออกมาเป็นค่าตัวเลขที่แสดง ถึงความเข้มของความรู้สึกในเรื่องนั้น

รูปแบบคำตอบและจำนวนมาตราของมาตรประเมินค่า

รูปแบบคำตอบ

มาตรประเมินค่า (Rating Scale) เป็นเครื่องมือที่สะดวกในการสร้างและการนำไปใช้ ดังนั้น จึงเป็นเครื่องมือที่นิยมใช้กันอย่างแพร่หลาย ทั้งในด้านการศึกษาหรือโครงการต่างๆ จึง ได้มีการพัฒนาปรับปรุงรูปแบบให้แตกต่างกันออกไป นักการศึกษา และนักวิจัยหลายท่านได้ พยายามวิเคราะห์ และจำแนกรูปแบบของมาตรประเมินค่าตามลักษณะที่ปรากฏ โดยมากมัก จำแนกตามลักษณะการตอบของมาตรประเมินค่าดังนั้นจึงกล่าวได้ว่ารูปแบบของมาตรก็คือ รูปแบบคำตอบนั่นเอง สรุปในรูปแบบต่างๆ ได้ ดังต่อไปนี้

กิลฟอร์ด (Guilford.1954 : 263 – 275) แบ่งมาตรประเมินค่าออกเป็น 5 รูปแบบ คือ

1. แบบตัวเลข (Numerical Scales) กำหนดตัวเลขพร้อมกับนิยามหรือคำบรรยายไว้ให้ผู้ประเมิน ดังตัวอย่างมาตรวัดที่ใช้ในการประเมินค่าของสีและกลิ่นต่อไปนี้

- | | |
|----|--|
| 10 | พอใจอย่างที่สุด (Most Pleasant Imaginable) |
| 9 | พอใจมากอย่างยิ่ง (Most Pleasant) |
| 8 | พอใจอย่างยิ่ง (Extremely Pleasant) |
| 7 | พอใจพอสมควร (Moderately Pleasant) |
| 6 | พอใจปานกลาง (Mildly Pleasant) |
| 5 | ไม่พอใจ (Indifferent) |
| 4 | ไม่พอใจปานกลาง (Mildly Unpleasant) |
| 3 | ไม่พอใจพอสมควร (Moderately Unpleasant) |
| 2 | ไม่พอใจมากอย่างยิ่ง (Most Unpleasant) |
| 1 | ไม่พอใจมากอย่างยิ่งที่สุด (Most Unpleasant Imaginable) |

ในบางครั้งมาตรวัดแบบตัวเลข อาจไม่มีตัวเลขให้ มีแต่เพียงข้อความ (Descriptive Cues) ดังนี้

- หนักมากที่สุด (Very Heavy)
- หนักมาก (Heavy)
- หนักปานกลาง (Medium)
- เบา (Light)
- เบาที่สุด (Very Light)

2. แบบกราฟฟิก (The Graphic Scale) เป็นแบบที่นิยมใช้กันอย่างกว้างขวางมากที่สุด มีได้หลายรูปแบบ เน้นการจัดข้อความเข้ากับเส้นตรง จะอยู่ในแนวนอน หรือแนวตั้งก็ได้ แต่ในแนวนอนจะเป็นที่นิยมมากกว่า เส้นตรงดังกล่าว สามารถแบ่งออกเป็นช่วงๆ ซึ่งสามารถแบ่งได้หลายช่วงหรืออาจต่อเนื่องกันตลอด ดังตัวอย่างต่อไปนี้

ในวงสนทนาคุณเป็นอย่างไร

--- ---	--- ---	--- ---	--- ---	--- ---
ช่วงพูด	พูดบ่อย	พูดเมื่อ จำเป็น	ชอบฟัง มากกว่า	ไม่เคย พูดเลย

เขาคิดได้ช้าหรือเร็ว

-----	-----	-----	-----	-----
ช้ามาก	ค่อยๆ คิด	ธรรมดา	ว่องไว	รวดเร็วมาก

3. แบบมาตรฐาน (Standard Scales) เป็นแบบที่ได้กำหนดค่าสเกล เป็นค่าประจำของชุดมาตรฐานที่จะใช้ในการเปรียบเทียบไว้ล่วงหน้า ผู้ประเมินจะต้องทำการเปรียบเทียบรายการที่ให้ประเมินกับชุดมาตรฐาน มาตรฐานที่จัดอยู่ในรูปแบบนี้ได้แก่วิธีจับคู่เหมือน (Portrait – Matching Scale) วิธีเปรียบเทียบคนต่อคน (Man – to – Man Scale) ซึ่งมาตรฐานแบบนี้ไม่เป็นที่นิยมเท่าที่ควร

4. การประเมินค่าโดยใช้แต้มสะสม (Rating by Cumulative Points) แตกต่างจากแบบอื่นที่วิธีการให้คะแนน วิธีการตรวจรายการ (Check – list) เป็นตัวอย่างหนึ่งของรูปแบบนี้ ได้มีผู้ใช้วิธีการตรวจรายการในการประเมินคุณลักษณะของเด็ก โดยกำหนดรายการที่มีทั้งคุณลักษณะที่ดีและไม่ดีมาให้เลือก เช่น ให้ความร่วมมือ ช่างคิด ดุร้าย เห็นแก่ตัว เป็นต้น ผู้ประเมินแต่ละคนจะเลือกทุกๆ คุณลักษณะที่ตรงกับเด็กคนนี้ คะแนนของเด็กแต่ละคนจะเป็นผลรวมของคะแนนทั้งหมด ซึ่งได้กำหนดให้ 1 สำหรับทุกลักษณะที่ดี และให้คะแนน - 1 สำหรับทุกลักษณะที่ไม่ดี นอกจากนี้แบบตรวจรายการอาจจะจัดอยู่ใน รูปแบบตัวเลือก (Multiple – Choice) ก็ได้ ดังตัวอย่าง

การให้ความร่วมมือกับผู้อื่น

----- กระตือรือร้น

----- ให้ความสนใจพอสมควร

----- ไม่ค่อยให้ความสนใจ

----- ไม่เต็มใจให้ความร่วมมือ

----- ไม่ยอมให้ความร่วมมือ

5. การประเมินค่าแบบบังคับเลือก (Forced –Choice Rating) เป็นแบบที่พัฒนาขึ้นเพื่อประเมินบุคคล ผู้ประเมินไม่ต้องบอกว่าผู้ถูกประเมินมีลักษณะใด มากน้อยเพียงใด แต่ให้เปรียบเทียบเป็นรายคู่ แล้วบอกว่ามีลักษณะใดมากกว่ากัน ดังตัวอย่าง

----- เลินเล่อ สะเพร่า

----- เอาใจใส่

----- ขะมักเขม้น

----- วางภูมิ หิวสูง

แต่ละข้อจะประกอบด้วยข้อความคู่ 2 คู่ ข้อความคู่แรกเป็นไปในทางบวก ส่วนข้อความคู่ที่เหลือจะเป็นไปในทางลบ การที่จัดรูปแบบดังกล่าวก็เพราะว่าในบางครั้งผู้ประเมินไม่ได้จุดมุ่งหมายที่จะเลือก ข้อความในทางบวก แต่ก็จะทำให้ลังเลที่จะเลือกข้อความในทางลบเพราะยังมีข้อความในทางบวกให้พิจารณาอีก วิธีประเมิน และเลือกอีกข้อความหนึ่งที่ไม่เหมาะสมหรือไม่ตรงกับลักษณะของผู้ถูกประเมิน

ทอลล์และฮอว์กินส์ (Tull and Hawkilns. 1993 : 373 – 389) ได้แบ่งมาตรประเมินค่าออกเป็น 2 ประเภทใหญ่ๆ ดังนี้ คือ

1. แบบไม่มีเกณฑ์เปรียบเทียบให้ (No comparative Rating Scales) แบบนี้ ผู้ตอบจะประเมินค่าสิ่งเร้านั้นตามความรู้สึกนึกคิดของตนเปรียบเทียบกับเกณฑ์อะไรก็ได้ ที่ผู้ตอบแต่ละคนจะเห็นว่าเหมาะสม ต่างคนต่างกำหนดเอง ไม่เป็นมาตรฐานเดียวกัน แบ่งเป็น 2 ชนิดได้แก่

1.1 มาตรประเมินค่าแบบกราฟฟิก (Graphic Rating scales) ชนิดนี้พยายามทำให้การประเมินมีลักษณะต่อเนื่อง โดยให้ผู้ตอบประเมินสิ่งเร้าด้วยการทำเครื่องหมายลงบนเส้นกราฟ มี 2 ลักษณะ คือ แบบเป็นเส้นกราฟตรงและมีเฉพาะตัวหนังสือบรรยายคุณลักษณะกำกับหัวท้ายเท่านั้น กับแบบเป็นเส้นกราฟตรง มีตัวหนังสือบรรยายคุณลักษณะกำกับไว้เป็นระยะๆ ดังตัวอย่าง

นักเรียนสามารถพุดจูงใจผู้อื่นได้ดีเพียงใด

ดีมาก _____ แย่มาก

นักเรียนสามารถพุดจูงใจผู้อื่นได้ดีเพียงใด

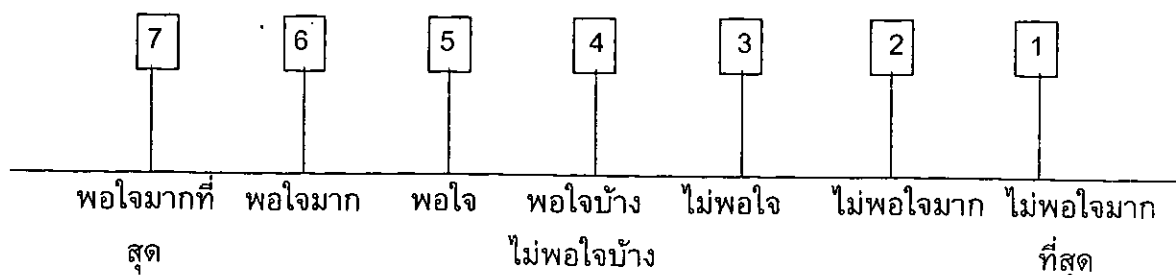
ดีที่สุด ดีมาก ไม่ดีไม่แย ไม่ดีเลย แย่มาก

มาตรประเมินค่าด้วยเส้นกราฟนี้ เส้นตรงที่กำหนดให้ทุกข้อจะต้องมีความยาวเท่ากัน และไม่มี การแบ่งช่วงบนเส้นไว้ให้ก่อน ผู้ตอบจะเลือกกากบาทลงบนเส้นกราฟ ณ จุดใดก็ได้ ตามแต่ผู้ตอบจะคาดคะเนเอง

1.2 มาตรประเมินค่าด้วยคำตอบจำกัด (Itemized Rating Scales) ชนิดนี้เป็นแบบกำหนดข้อความ คำ ตัวเลข หรือสัญลักษณ์ไว้เป็นคำตอบให้เลือกแน่นอน มี 3 ลักษณะได้แก่

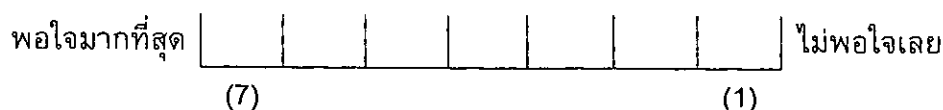
1.2.1 แบบกำหนดคำตอบเป็นข้อความลักษณะนี้จะเลือกคำหรือข้อความที่มีลักษณะการเปรียบเทียบกันมาเรียงตามลำดับมากน้อยไว้ให้เลือกตอบ ดังตัวอย่าง

ท่านมีความพอใจต่อความเป็นอยู่ในชีวิตของท่านเพียงใด

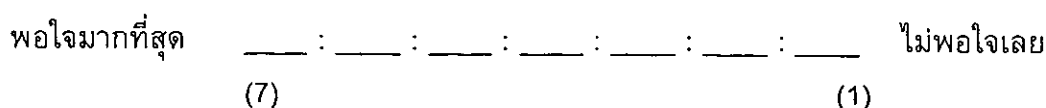


1.2.2 แบบกำหนดคำตอบเป็นเส้นกราฟ ลักษณะนี้กำหนดเป็นกราฟเส้น มีคำบรรยายกำกับหัวท้าย แต่เส้นกราฟนั้นแบ่งเป็นคำตอบกำหนดไว้แน่นอนและจะมีตัวเลขกำกับไว้ด้วยก็ได้ ดังตัวอย่าง

ท่านมีความพอใจต่อความเป็นอยู่ในชีวิตของท่านเพียงใด

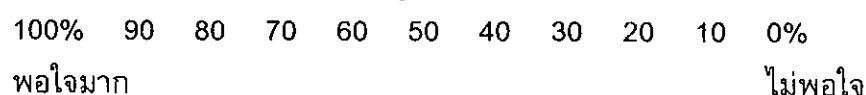


สำหรับเส้นกราฟนี้อาจจะใช้เส้นกราฟที่ไม่ต่อกันก็ได้ ดังตัวอย่าง



1.2.3 แบบกำหนดคำตอบเป็นร้อยละ ลักษณะนี้กำหนดปริมาณคำตอบไว้เป็นร้อยละโดยประมาณ โดยแบ่ง 100% เป็นช่วงๆ เท่าๆ กัน และมีข้อความหรือคำกำกับหัวท้าย ดังตัวอย่าง

ท่านมีความพอใจต่อความเป็นอยู่ในชีวิตของท่านเพียงใด



2. แบบมีเกณฑ์เปรียบเทียบ (Comparative Rating Scales) แบบนี้จะกำหนดเกณฑ์ไว้ ให้เปรียบเทียบ ผู้ตอบทุกคนต้องใช้เกณฑ์เปรียบเทียบเดียวกัน แบ่งเป็น 4 ชนิด ได้แก่

2.1 มาตรฐานค่าแบบกราฟิก และคำตอบจำกัดในลักษณะเปรียบเทียบ (Graphic and Itemized Comparative Rating Scales) ชนิดนี้ก็คือแบบที่ไม่มีเกณฑ์เปรียบเทียบให้ตนเอง เพียงแต่เพิ่มเกณฑ์ที่ใช้เปรียบเทียบมาให้ด้วยเท่านั้น ดังตัวอย่าง

กิจกรรมวันไหว้ครูเป็นอย่างไร เมื่อเทียบกับกิจกรรมอื่นๆ ที่โรงเรียนจัด
 ชอบมากกว่ามากๆ _____ ชอบน้อยกว่ามากๆ
 นักเรียนชอบกิจกรรมวันไหว้ครูมากเพียงใด เมื่อเทียบกับกิจกรรมกีฬา

 ชอบมากกว่า ชอบมากกว่า ชอบพอๆ กัน ชอบน้อยกว่า ชอบน้อยกว่า
 มากๆ _____ มากๆ

2.2 มาตรฐานประเมินค่าแบบเปรียบเทียบรายคู่ (Paired Comparisons) แบบนี้เป็นวิธีการ เปรียบเทียบของ 2 สิ่งในเวลาเดียวกัน และผู้ตอบจะต้องเลือกเห็นด้วยกับข้อความใดข้อความหนึ่ง ดังนั้น คำตอบที่กำหนดให้จะต้องเป็นไปตามจริง เช่น

จากข้อความ ก กับ ข นักเรียนเห็นด้วยกับข้อใดมากกว่า

ก. วิชาคณิตศาสตร์มีความสำคัญมากกว่าวิชาสังคมศาสตร์

ข. วิชาสังคมศึกษามีความสำคัญมากกว่าวิชาคณิตศาสตร์

ในบางกรณี หากต้องเปรียบเทียบมากกว่า 2 สิ่ง เช่น ต้องการเปรียบเทียบถึง 5 วิชา ($n = 5$) ก็จะมีข้อความให้เลือกถึง 10 คู่ ดังนั้น การเปรียบเทียบรายคู่จึงมีข้อจำกัดว่าจะเปรียบเทียบแค่ 2 – 3 สิ่งเท่านั้น

2.3 มาตรฐานประเมินค่าแบบจัดอันดับ (Rank Order Rating Scales) แบบนี้คล้ายกับการเปรียบเทียบรายคู่ เพียงแต่ให้เปรียบเทียบพร้อมกันทั้งหมด เช่น

ในการซื้อรถยนต์ท่านจะคำนึงถึงสิ่งใดบ้าง เรียงตามลำดับความสำคัญด้วยการให้ หมายเลข 1 – 5

ก. ความชอบ

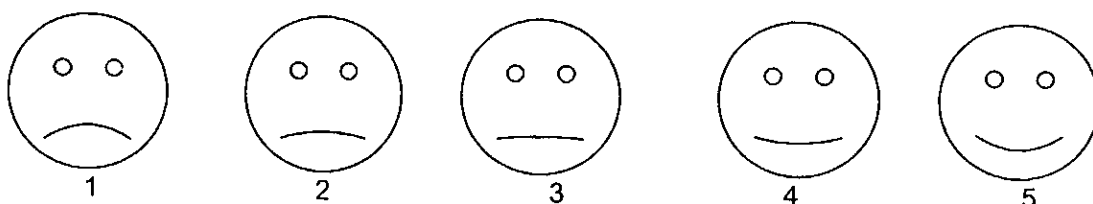
ข. ความประหยัด

ค. ความปลอดภัย

ง. ความสบาย

จ. รูปแบบ

2.4 มาตรฐานประเมินค่าแบบรอยยิ้ม (Smiling Face Scales) แบบนี้จะเหมาะสมสำหรับเด็กอายุประมาณ 5 ขวบ โดยจะมีรูปใบหน้าแสดงอาการต่างๆ กำหนดให้ จากนั้นครูจะอธิบายวิธีปฏิบัติและถามคำถาม ดังตัวอย่าง



“บอกครูว่า นักเรียนชอบมาโรงเรียนมากแค่ไหน โดยชี้ที่รูปภาพ ถ้านักเรียนไม่ชอบมาโรงเรียนเลยให้ชี้รูปที่ 1 ถ้าชอบมาโรงเรียนมากๆ ให้ชี้รูปที่ 5 นักเรียนชอบมาโรงเรียนมากแค่ไหน”

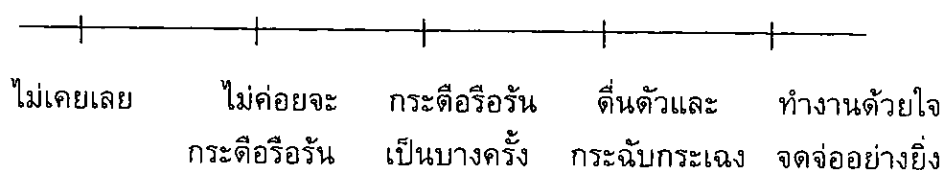
รัมเมล (Rummel, 1964 : 201 – 202) ได้กล่าวถึงมาตรประเมินค่าว่า ที่เป็นที่นิยมใช้มีหลายรูปแบบ พร้อมทั้งยกตัวอย่างไว้ 3 รูปแบบ ดังต่อไปนี้

1. มาตรประเมินค่าแบบตัวเลข (Numerical Rating Scales) ให้ตัดสินลักษณะออกมาเป็น ตัวเลข เช่น ให้ประเมินบุคคลตามข้อกระทงต่อไปนี้ โดยวงกลมล้อมรอบ 1 เมื่อไม่พอใจ 2 เมื่ออยู่ในระดับต่ำ 3 เมื่ออยู่ในระดับปานกลาง 4 เมื่อดีกว่าปานกลาง และ 5 เมื่ออยู่ในระดับที่ดี

มีบุคลิกดึงดูดความสนใจ	1	2	3	4	5
พิถีพิถันกับงานที่ทำ	1	2	3	4	5

2. มาตรประเมินค่าแบบกราฟิก (Graphic Rating Scales) ประกอบด้วยเส้นตรงสำหรับให้ ผู้ประเมินค่าทำเครื่องหมายเพื่อแสดงตำแหน่งที่เป็นจริงที่สุดเกี่ยวกับการถูกประเมิน

ความกระตือรือร้น



3. มาตรประเมินค่าแบบบรรยาย (Descriptive Rating Scales) มักเป็นคำบรรยายแสดงระดับต่างๆ ของบุคลิกลักษณะ ดังตัวอย่างการประเมินพฤติกรรม โดยให้เลือกคำหรือข้อความที่บรรยายถึงพฤติกรรมของนักเรียนของนักเรียนตามที่สังเกตได้

นักเรียนคนนี้แข็งแรงหรือไม่

- แข็งแรงและคล่องแคล่ว
- ไม่ใคร่เหนื่อย
- มีความอดทน
- เจื่องหงอย เหนื่อยง่าย

นอกจากนี้วอคเคิล (Vockell. 1995 :142-143) ได้แบ่งรูปแบบของมาตรประเมินค่าไว้ ดังนี้

แบบที่ 1 ให้นำค่าตัวเลขที่ตรงกับระดับพฤติกรรมของนักเรียนไปใส่ข้างหน้าข้อความที่เป็นพฤติกรรมนั้น เช่น

- 5 หมายถึง ดีมาก
- 4 หมายถึง ดี
- 3 หมายถึง ปานกลาง
- 2 หมายถึง พอใช้
- 1 หมายถึง ยังใช้ไม่ได้

-----ยอมรับฟังความคิดเห็นของคนอื่น

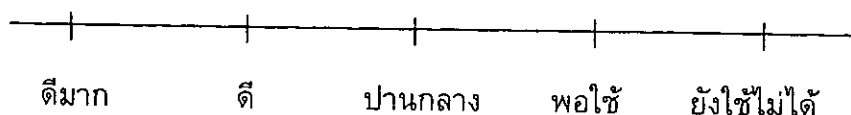
-----ฟังการบรรยายอย่างตั้งใจ

แบบที่ 2 เหมือนกับแบบแรก เพียงแต่ให้ผู้ประเมินเลือกตัวเลขที่กำหนด ดังนี้

ยอมรับฟังความคิดเห็นของคนอื่น	1 2 3 4 5
ฟังคำบรรยายอย่างตั้งใจ	1 2 3 4 5

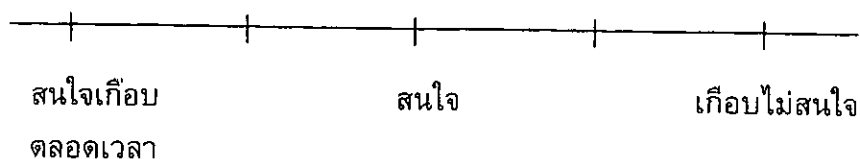
แบบที่ 3 เป็นแบบกราฟฟิก โดยให้วงกลมล้อมรอบข้อความที่ตรงกับพฤติกรรม เช่น

นักเรียนตั้งใจฟังคำบรรยายเพียงไร



แบบที่ 4 แบบกราฟฟิก เช่นเดียวกันเพียงแต่ให้ผู้ประเมินทำเครื่องหมายถูกลงบนเส้นตรงในตำแหน่งที่ตรงกับระดับพฤติกรรมที่สังเกตได้ เช่น

นักเรียนตั้งใจฟังคำบรรยายเพียงไร



เซ็ง แวร์ และพาเรตแวจ (ประภัสสร สุขชื่น.2540 ; 15 ; Tzeng Ware and Bharadwaj. 1991 : *Education and Psychological Measurement.*) ได้แบ่งมาตรประเมินค่าไว้ 2 รูปแบบ ดังนี้

1. มาตรวัดแบบขั้วเดียว (Unipolar) ในแต่ละข้อจะมีข้อคำถามที่เป็นข้อความเดียว มีสเกลให้ผู้ตอบทำเครื่องหมายในช่องที่ตรงกับระดับพฤติกรรมที่เป็นจริง

ข้าพเจ้าชอบการ สังสรรค์	7	6	5	4	3	2	1
----------------------------	---	---	---	---	---	---	---

2. มาตรวัดแบบสองขั้ว (Bipolar) ในแต่ละข้อจะมีข้อคำถามอยู่สองข้อความซึ่งมีลักษณะที่ตรงข้ามกันเป็นคู่ ๆ ให้ผู้ตอบเลือกตอบลงบนสเกลทางใดทางหนึ่งที่ตรงกับระดับพฤติกรรมที่เป็นจริงมากที่สุด

ข้าพเจ้าชอบการ สังสรรค์	3	2	1	0	-3	-2	-1	ข้าพเจ้าชอบอยู่คนเดียว
----------------------------	---	---	---	---	----	----	----	------------------------

อุทุมพร จามรมาน (2538 : 69) ได้แบ่งมาตรประเมินค่าออกเป็น 2 แบบ ดังต่อไปนี้

1. แบบขั้วเดียว (Unipolar) คือ ในแต่ละข้อมีข้อความทางเดียวเป็นบวกหรือลบ อย่างใดอย่างหนึ่ง เช่นมากที่สุด มาก ปานกลาง (มาตรวัดขั้วเดียวทางบวก) หรือปานกลาง น้อย น้อยที่สุด (มาตรวัดขั้วเดียวทางลบ)

2. แบบสองขั้ว (Bipolar) คือ ในแต่ละข้อมีข้อความที่มีลักษณะตรงข้ามกันเป็นคู่ ๆ ซึ่งมีน้ำหนักเท่ากัน 2 ข้าง เช่น มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด

อนันต์ ศรีโสภ (2521 : 213 – 215) ได้แบ่งมาตรประเมินค่าออกเป็น 5 รูปแบบ ดังต่อไปนี้

1. แบบตัวเลข (Numerical Rating Scales) นับว่าเป็นการ rating ที่ง่ายที่สุด ผู้ประเมินเพียงแต่กาเครื่องหมายลงบนตัวเลขที่แสดงระดับ หรือคุณภาพของการกระทำ หรือลักษณะนิสัย เช่น (trait) ในขณะนั้นโดยมีสเกลตั้งแต่ 1 ถึง 5 แต่ไม่ควรเกิน 10 และควรบอกความหมายของสเกลการวัดแต่ละค่าด้วย

จำนวนมาตรา

จำนวนมาตรามีตั้งแต่ 2 ช่วง ขึ้นไป ขึ้นอยู่กับว่าผู้สร้างต้องการให้ได้คำตอบจากผู้ตอบ
มาตราละเอียดหรือหยาบมากน้อยเพียงใด การตัดสินใจใช้เพียง 2 ช่วง ก็หมายความว่า ผู้สร้าง
ไม่ต้องการให้ผู้ตอบจำแนกแยกแยะมากนัก เช่น

ท่านพอใจการสอนของอาจารย์ ก ใช่ ไม่ใช่

ในกรณีที่ผู้สร้างต้องการคำตอบละเอียดกว่า อาจทำได้ดังนี้

ท่านพอใจการสอนของอาจารย์ ก มาก ปานกลาง น้อย

หรือถ้าต้องการให้ละเอียดมากขึ้นอาจเป็น

ท่านพอใจการสอนของอาจารย์ ก

มากที่สุด

มากรองลงมา

มากพอสมควร

มาก

ปานกลาง

น้อย

น้อยมาก

น้อยที่สุด

การกำหนดจำนวนช่วงของมาตราประเมินค่าว่าควรจะมีกี่ช่อง และควรจะเป็นเลขคู่
หรือเลขคี่เป็นอีกประเด็นหนึ่งที่มีผู้สงสัยมาก โดยทั่วไปการกำหนดจำนวนช่วงขึ้นอยู่กับ
ความต้องการของผู้สร้างว่า จะให้มีการวัดทางจิตวิทยานั้นๆ ละเอียดมากเพียงใด ถ้าผู้สร้าง
ต้องการคำตอบที่หยาบ ๆ เช่น ใช่-ไม่ใช่ เห็นด้วย-ไม่เห็นด้วย ก็มี 2 ช่วง แต่ถ้าต้องการ
คำตอบที่ละเอียด ก็อาจเพิ่มจำนวนช่วงเป็นเลขคู่ เช่น 4 ช่วง ดังตัวอย่าง

เห็นด้วยมาก เห็นด้วย ไม่เห็นด้วย ไม่เห็นด้วยมาก
4 3 2 1

การกำหนดช่วงตรงกลาง

มาตรฐานค่าที่มีการกำหนดช่วงตรงกลาง หมายถึง จะต้องมีการกำหนดช่วงที่อยู่ ขวาและซ้ายของช่วงตรงกลาง ดังตัวอย่าง

เห็นด้วย	ตัดสินใจไม่ได้	ไม่เห็นด้วย
3	2	1
เห็นด้วย		ไม่เห็นด้วย
2		1

ประโยชน์ของมาตรประเมินค่า

มาตรประเมินค่า (Rating Scale) ใช้รับบุคคลให้ตอบสนองโดยแสดงระดับความรู้สึก หรือ เจตคติที่มีต่อสถานการณ์ เรื่องราว คุณค่า หรือความต้องการ แล้วมีการตรวจให้คะแนน แทนระดับความรู้สึก หรือเจตคตินั้น ดังนั้นมาตรประเมินค่าจึงสามารถนำมาใช้ ในการประเมิน ผลการเรียนรู้และพัฒนาการต่าง ๆ ได้หลายด้าน ดังนี้

1. ใช้ประเมินกระบวนการ เช่น (Procedure Evaluation) ใช้ในการประเมินทักษะ ในการพูด การจัดเตรียมเครื่องมือในการทดลอง การทำงานเป็นกลุ่มในรายวิชาต่าง ๆ เป็นต้น
2. ใช้ในการประเมินผลผลิต (Product Evaluation) เช่น การคัดลายมือ การวาดภาพระบายสี การทำรายงาน การทำงานประดิษฐ์ เป็นต้น
3. ใช้ในการประเมินพัฒนาการทางบุคลิกภาพและทางสังคม (Evaluating Personal-Social Development) เช่น ความคิดสร้างสรรค์ ความประหยัด ความเป็นผู้นำ ความมั่นคงทางอารมณ์ ความรับผิดชอบ และทัศนคติ เป็นต้น

ความคลาดเคลื่อนจากการใช้มาตรประเมินค่า

การใช้มาตรประเมินค่านั้นมีแหล่งที่ก่อให้เกิดความคลาดเคลื่อนใหญ่ๆ 3 แหล่ง คือ

1. ความบกพร่องของเครื่องมือ ซึ่งข้อบกพร่องที่สำคัญของมาตรประเมินค่ามีอยู่ 2 ส่วน คือ ส่วนของคุณลักษณะที่เป็นชุดของสิ่งเร้า และส่วนของคำตอบที่เป็นสเกลแสดงระดับ ความเข้มของพฤติกรรม ซึ่งความบกพร่องในส่วนของสิ่งเร้าเกิดขึ้น อาจเป็นเพราะ
 - 1) ลักษณะที่เลือกมาไม่มีความสำคัญหรือไม่น่าสนใจ
 - 2) ขาดการอธิบายคุณลักษณะของสิ่งที่เลือกมาให้ประเมินหรืออธิบายไม่ชัดเจน ทำให้เกิดความไม่เข้าใจไม่ตรงกัน ทำให้ผู้ตอบแบบหรือผู้ประเมินไม่สนใจและไม่เต็มใจตอบ หรือเต็มใจก็ไม่สามารถตอบได้ตรงตามความเป็นจริงได้จึงส่งผลให้การประเมินขาดความเชื่อมั่น

ส่วนข้อบกพร่องที่เกิดจากส่วนของคำตอบนั้นอาจเนื่องมาจาก

- 1) จำนวนช่วงหรือสเกลไม่เหมาะสม
- 2) การให้ความหมายของสเกลไม่ชัดเจน
- 3) การให้ความหมายของแต่ละสเกลคาบเกี่ยวกันจนไม่สามารถจำแนกความแตกต่างได้

4) ลักษณะการให้ความหมายของสเกลไม่สอดคล้องกับลักษณะที่เป็นสิ่งเร้า

2. ความบกพร่องที่เกิดจากกระบวนการในการประเมิน ซึ่งข้อบกพร่องที่มักจะพบคือ คุณลักษณะหรือสิ่งที่ต้องการให้ประเมินมีจำนวนมากเกินไป ผู้ประเมินไม่สนใจและเอาใจใส่ ข้อบกพร่องที่สำคัญ อีกประการหนึ่งคือ ไม่สามารถหาผู้ประเมินหลายคนมาทำการประเมิน นักเรียนคนเดียวหรือห้องเดียวได้ ซึ่งการมีผู้ประเมินหลายคนจะเป็นวิธีที่ช่วยลดความคลาดเคลื่อนที่เกิดจากแต่ละบุคคลได้

3. ความคลาดเคลื่อนที่เกิดจากตัวผู้ประเมิน ซึ่งมีความคลาดเคลื่อนหลายประการที่พึงระมัดระวัง ได้แก่ ความลำเอียงหรือความมีอคติของแต่ละบุคคล (Personal Bias) ผลจากการปฏิสัมพันธ์มาก่อน (Halo effect) ความคลาดเคลื่อนแบบใช้ตรรกผิด (Logical error) และความคลาดเคลื่อนแบบที่เกินหรือขาดจากความเป็นจริง (Error of Leniency)

จากการที่แหล่งความคลาดเคลื่อนที่สำคัญมี 3 แหล่งดังกล่าวแล้วนั้น พบว่าแหล่งที่มีปัญหามากที่สุด ได้แก่ ความคลาดเคลื่อนที่เกิดจากตัวผู้ประเมินหรือผู้ตอบ ซึ่งความคลาดเคลื่อนที่พบมากในการใช้มาตรประเมินค่า ได้แก่

1. ความลำเอียงหรือความมีอคติส่วนบุคคลของผู้ประเมิน (Personal Bias) เป็นความคลาดเคลื่อนที่เกิดขึ้นเฉพาะลักษณะนิสัยประจำตัวของผู้ประเมิน ที่มักมีรูปแบบการประเมินที่เฉพาะ (fixed pattern) ซึ่งมี 3 ชนิด คือ

1) ความคลาดเคลื่อนจากการประเมินที่ให้ค่าสูงกว่าความเป็นจริง (Generosity Error) เป็นความคลาดเคลื่อนที่เกิดเมื่อผู้ประเมินให้ค่าคุณลักษณะต่างๆ ของทุกคนในระดับที่ค่อนข้างสูงเพียง ด้านเดียว ทำให้ค่าที่ได้สูงกว่าความเป็นจริง

2) ความคลาดเคลื่อนที่เกิดจากการประเมินค่าต่ำกว่าความเป็นจริง (Severity Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมินให้ค่าคุณลักษณะต่าง ๆ ของทุกคนในระดับค่อนข้างต่ำเพียงด้านเดียวเป็นผลให้ค่าที่ประเมินได้ต่ำกว่าที่ควรจะเป็น

3) ความคลาดเคลื่อนแบบโน้มเข้าสู่ส่วนกลาง (Central Tendency Error) เกิดขึ้นเมื่อผู้ประเมินพยายามหลีกเลี่ยงการให้คะแนนที่ค่อนข้างสูงและค่อนข้างต่ำมาเป็นให้คะแนนในระดับกลางๆ ทำให้คะแนนที่ได้สูงหรือต่ำกว่าความเป็นจริง ซึ่งสาเหตุของความคลาดเคลื่อนอาจเป็นเพราะ ผู้ประเมินอาจไม่ทราบว่สิ่งที่เขาประเมินมีพิสัยอยู่ในทางสูงหรือต่ำ ผู้ประเมินอาจไม่มีความเจียวคมในการจำแนก ความแตกต่างของสิ่งที่ประเมิน อาจเป็นเพราะบุคลิกของ

ผู้ประเมินเป็นคนไม่มั่นคง ไม่กล้าตัดสินใจที่จะให้ค่าสูงหรือค่าต่ำ ผู้ประเมินอาจรู้สึกว่าเป็นการไม่ยุติธรรม เพราะเกรงว่าจะเป็นการทำร้ายจิตใจกัน เป็นต้น

การประเมินค่าโดยให้ค่าในระดับหนึ่งเพียงระดับเดียว จะส่งผลให้เกิดสิ่งไม่พึงประสงค์ 2 ประการ คือ

ประการแรก ทำให้ค่าที่ได้จากการประเมินขาดความตรงจากสภาพที่แท้จริง ทั้งนี้ เพราะเป็นการให้คะแนนตามบุคคลมากกว่าเป็นคะแนนของคุณลักษณะนั้นๆ

ประการที่สอง การที่ผู้ประเมินมักให้ค่าเฉพาะระดับใดระดับหนึ่ง เป็นการจำกัดขอบเขตของการประเมิน ทำให้อ่านาจจำแนกค่า ความเชื่อมั่นก็ย่อมต่ำด้วย

2. ผลจากการมีปฏิสัมพันธ์มาก่อน (Halo Effect) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ถูกประเมินมีอิทธิพลต่อผู้ประเมิน คือ จะให้ค่ามากหรือน้อยขึ้นอยู่กับความพอใจของผู้ประเมินต่อสิ่งนั้น ซึ่งความคลาดเคลื่อนแบบนี้จะเกิดเฉพาะผู้ถูกประเมินบางคนเท่านั้น

3. ความคลาดเคลื่อนแบบใช้ตรรกผิด (Logical Error) เกิดขึ้นเมื่อผู้ประเมินนำตัวแปรอื่นที่ตนเองคิดว่ามีความสัมพันธ์กับตัวแปรที่ประเมิน มามีส่วนในการประเมินด้วย ความคลาดเคลื่อนประเภทนี้ มิได้เกิดจากความลำเอียง แต่เกิดจากการรับรู้มาก่อนของผู้ประเมิน

4. ความคลาดเคลื่อนที่เกินหรือขาดจากความเป็นจริง (Error of Leniency) เกิดจากความพยายามของผู้ประเมินที่จะให้ในด้านดีเกินไป (Easy Rate) หรือในทางตรงกันข้ามผู้ประเมินพยายามประเมินในด้านต่ำกว่าความเป็นจริง (Hard Rate)

การปรับปรุงและแก้ไขการใช้มาตรฐานประเมินค่าให้มีประสิทธิภาพ

เนื่องจากการใช้มาตรฐานประเมินค่าก่อให้เกิดความคลาดเคลื่อนดังกล่าวแล้วนั้น จำเป็นต้องหาวิธีการลดความคลาดเคลื่อนให้น้อยลง ซึ่งวิธีการที่เหมาะสมมีสองวิธี คือ การออกแบบมาตรฐานประเมินค่าให้ถูกต้องเหมาะสม และมีวิธีการใช้อย่างถูกต้องและเหมาะสมด้วยเพื่อให้มีความเชื่อถือได้ พอสรุปสาระสำคัญได้ดังนี้

การสร้างมาตรฐานประเมินค่า

1. ส่วนของชุดสิ่งเร้าที่เป็นลักษณะที่จะให้ประเมินหรือพฤติกรรมที่เลือกมาควรมีลักษณะดังนี้

- 1) เป็นสิ่งที่มีความสำคัญตรงกับสิ่งที่ต้องการวัดมากที่สุด
- 2) ต้องเป็นอิสระไม่ก้ำกึ่งกับลักษณะอื่น
- 3) เป็นสิ่งที่สามารถสังเกตเห็นได้โดยตรง
- 4) มีความเป็นเอกภาพ ไม่ใช่องค์ประกอบของลักษณะที่มีความสัมพันธ์กันต่ำ
- 5) คุณลักษณะที่จะให้ประเมินไม่ควรมีมากเกินไป

2. ส่วนของคำตอบที่ให้ความหมายระดับพฤติกรรมจะต้องสร้าง ดังนี้

1) ให้คำนิยามหรือบรรยายระดับพฤติกรรมของแต่ละลักษณะที่จะประเมินให้ชัดเจน เพื่อให้เกิดความชัดเจนเป็นที่เข้าใจตรงกันของผู้ประเมิน จะทำให้สามารถตัดสินใจได้อย่างเป็นปรนัยและมีความเที่ยง

2) สิ่งที่อธิบายระดับพฤติกรรมจะต้องมีความต่อเนื่องกันตามลำดับทุกสเกล

3) คำอธิบายสเกลแต่ละพฤติกรรมจะต้องสอดคล้องกับแต่ละลักษณะ

4) ควรเป็นคำอธิบายที่เข้าใจได้ง่ายและตรงกันสามารถมองเห็นเป็นรูปธรรมมากกว่าที่จะใช้คำวิเศษณ์ง่าย ๆ ทั่ว ๆ ไป

5) การอธิบายในทุกสเกลจะต้องเป็นเรื่องเดียวกันและมีความต่อเนื่อง

6) จำนวนช่วงหรือสเกลที่จะประเมินควรสร้างให้เหมาะสม ดังนี้

ก) ปกติมักแบ่งช่วงออกเป็น 3 – 7 ช่วง การแบ่งช่วงยิ่งละเอียดมากยิ่งขึ้นแสดงความแตกต่างได้มาก แต่ไม่ควรมากเกินไปจนแยกความแตกต่างไม่ได้

ข) มีการนิยามคำตอบที่ต่อเนื่องของสเกลอย่างชัดเจน

ค) ควรมีการสลับช่วงหรือลำดับพฤติกรรมจากมากไปหาน้อย และจากน้อยไปหามาก เพื่อที่จะแก้ปัญหาการประเมินโดยอัตโนมัติ

จากรูปแบบต่าง ๆ ของมาตรประเมินค่าตามที่กล่าวมาจะเห็นว่ารูปแบบที่นิยมใช้ส่วนใหญ่ ได้แก่ มาตรประเมินค่าแบบชั่วเดียว แบบสองชั่ว แบบเติมตัวเลข แบบคำตอบจำกัด แบบบรรยาย แบบกราฟฟิค และแบบตัวเลข โดยแต่ละรูปแบบสามารถกำหนดจำนวนมาตราตามความต้องการของผู้สร้างอาจกำหนดเป็นจำนวนคู่ หรือ อาจกำหนดเป็นจำนวนคี่ เช่น 5 ช่วงก็ได้ จากการศึกษาพบว่าจำนวนช่วงที่เหมาะสมและ ลดความคลาดเคลื่อนได้ คือ จำนวน 3-7 ช่วง แต่โดยปกตินิยมกำหนด 3 ถึง 8 ช่วง โดยการใช้มาตรประเมินค่าก่อให้เกิดความคลาดเคลื่อน 3 แหล่ง คือ 1). ความบกพร่องของเครื่องมือ 2). ความบกพร่องที่เกิดจากกระบวนการในการประเมิน 3). ความบกพร่องที่เกิดจากตัวผู้ประเมิน โดยเฉพาะความบกพร่องที่เกิดจากตัวผู้ประเมินนั้นเป็นแหล่งที่ก่อให้เกิดความคลาดเคลื่อนหลายประเภท ดังนั้นในการสร้างมาตรประเมินค่าจำเป็นต้องหาวิธีการลดความคลาดเคลื่อนให้น้อยลง ซึ่งวิธีการที่เหมาะสมมีสองวิธี คือ การออกแบบมาตรประเมินค่าให้ถูกต้องเหมาะสม เช่น ในส่วนของลักษณะที่จะให้ประเมินต้องมีความชัดเจนเป็นเอกภาพสังเกตได้ และมีวิธีการใช้อย่างถูกต้องและเหมาะสมด้วยเพื่อให้มาตรประเมินค่ามีความเชื่อถือได้

2. บุคลิกภาพการทำให้ตนเองเสียเปรียบ

2.1 ความหมาย

เชดคัทดี โฆวาลินธุ์ (2520 : 62) ให้ความหมายว่า บุคลิกภาพเป็นลักษณะนิสัย (traits) ที่รวมกันเป็นแบบฉบับเฉพาะตัวของแต่ละบุคคล และเป็นสิ่งที่ทำให้เห็นความแตกต่างระหว่างบุคคล ซึ่งพิจารณาได้จากรูปแบบพฤติกรรมของบุคคลนั้นที่แสดงออกหรือตอบสนองต่อสิ่งแวดล้อม

มิเชล (Mischel. 1976 : 2) ให้ความหมายของบุคลิกภาพว่าเป็นลักษณะพิเศษแบบแผนพฤติกรรม (รวมทั้งความคิดและอารมณ์) ซึ่งเป็นลักษณะนิสัยประจำตัวของบุคคลที่สามารถปรับตัวให้เข้ากับเหตุการณ์หรือสถานการณ์

กัทธรี (Guthrie. 1944 : 58) ให้ความหมายว่า คือลักษณะนิสัยและระบบลักษณะนิสัยที่มีความสำคัญต่อสังคม อันมีสภาพคงตัวและยากที่จะเปลี่ยนแปลง

จากความหมายดังกล่าวข้างต้น สรุปได้ว่า บุคลิกภาพ คือ เอกลักษณะเฉพาะตัวของแต่ละบุคคล อาจเปลี่ยนแปลงได้บ้าง แต่ต้องใช้เวลา

2.2 การทำให้ตนเองเสียเปรียบ (Self – handicapping) (เปรมวดี กาญจนวีระ, 2541 : 3 – 14)

การให้ตนเองเสียเปรียบเป็นกลไกในการอธิบายถึงสาเหตุสำหรับพฤติกรรมที่ไม่ประสบความสำเร็จ มีการนำมาใช้เป็น 2 รูปแบบ คือ

1. การกระทำพฤติกรรมเพื่อให้ตนเองเสียเปรียบ (Behavioral Self – handicapping) เป็นการกระทำพฤติกรรมที่คาดว่าจะทำให้เกิดความสำเร็จได้ยากขึ้น เพื่อใช้เป็นสาเหตุในการอธิบายที่สนับสนุนตนเองเมื่อเกิดความล้มเหลว

เบอร์กลาส และ โจนส์ (Berglas and Jones.) เป็นผู้เสนอแนวคิดเรื่องการทำให้ตนเองเสียเปรียบ (Self – handicapping strategies) ในปี ค.ศ. 1978 โดยให้ความหมายของการทำให้ตนเองเสียเปรียบว่า เป็นการกระทำใดๆ ที่ได้จัดเตรียมไว้ล่วงหน้าก่อนที่จะมีการแสดงออก หากการกระทำนั้นประสบความสำเร็จล้มเหลว บุคคลจะหาข้ออ้างโดยยกสาเหตุว่ามาจากภายนอก แต่ถ้าประสบความสำเร็จก็จะยกสาเหตุว่า มาจากสาเหตุภายใน นั่นคือ ยอมรับว่าความสำเร็จนั้นมาจากความสามารถของตนเอง (Berglas & Jones. 1978)

โดยทฤษฎีแล้วบุคคลที่ทำให้ตนเองเสียเปรียบเป็นผู้ไม่แน่ใจระดับความสามารถที่แท้จริงของตนเอง ผู้ที่มีความนับถือตนเองต่ำ เป็นผู้ที่มีแนวโน้มว่าจะใช้วิธีการทำให้ตนเองเสียเปรียบ บุคคลที่มักใช้ วิธีการนี้ไม่ได้ตั้งใจเตรียมการเพื่อให้เกิดความล้มเหลว แต่พวกเขายินดีที่จะยอมรับความล้มเหลวที่จะเกิดขึ้น และถ้าประสบความสำเร็จก็จะเพิ่มคุณค่าต่อความนับถือตนเอง

โดยสรุปแล้ว สามารถกล่าวได้ว่า การทำให้ตนเองเสียเปรียบเป็นการยับยั้งให้ผู้อื่นเชื่อว่ามีปัจจัยที่ทำให้การกระทำของบุคคลนั้นตกอยู่ในสถานการณ์ที่จะได้รับความล้มเหลว

2. การกล่าวอ้างข้อเสียเปรียบของตนเอง หรือการใช้คำพูดอ้างถึงข้อเสียเปรียบ (Self – reported Handicaps) เป็นการกล่าวอ้างถึงความอ่อนแอหรือปัญหาที่เป็นสาเหตุทำให้บุคคลประสบความล้มเหลวในงาน

ซินเดอร์ และ สมิธ (Synder & Smith. 1982) ได้อภิปรายการใช้ลักษณะอาการของโรค เป็นวิธีการทำให้ตนเองเสียเปรียบ ความเป็นไปได้ประสบความล้มเหลว เป็นสิ่งคุกคามทำให้บุคคลเกิดความกลัว บางครั้งจึงแสดงอาการทางร่างกายหรือจิต เพื่อเป็นข้ออ้างว่าความล้มเหลวเป็นผลจากลักษณะอาการเหล่านี้ มากกว่าจะอ้างว่ามาจากการไร้ความสามารถ

ซึ่งสอดคล้องกับที่ ซินเดอร์ , สมิธ และเพอร์กินส์ (Smith, Snyder & Perkins. 1983) ได้ทำการศึกษา โดยตั้งสมมติฐานว่าผู้ที่วิตกกังวลเกี่ยวกับสุขภาพของตนเองมากเกินไป มีแนวโน้มมากกว่าคนอื่นๆ ที่จะรายงานว่าการกระทำของพวกเขาได้รับอุปสรรคจากอาการของโรคผลการทดลองเป็นไปตามสมมติฐานที่ตั้งไว้ว่า ผู้รับการทดลองที่วิตกกังวลเกี่ยวกับสุขภาพของตนเองมากเกินไป จะกล่าวอ้างถึงปัญหาสุขภาพที่เป็นสาเหตุทำให้เกิดความล้มเหลวในการทำแบบทดสอบวัดความสามารถทางสติปัญญา

จากงานวิจัยพฤติกรรมที่บ่งบอกว่าเป็นการกล่าวอ้างความเสียเปรียบของตนเอง (Self – reported Handicaps) แสดงถึงลักษณะเฉพาะของเรื่องการอ้างสาเหตุสนับสนุนตนเอง (Self – serving Attribution) ซึ่งเป็นแนวโน้มที่บุคคลจะสร้างเหตุผลที่จะสนับสนุนตนเองเมื่อประสบผลสำเร็จ แต่ปฏิเสธมันเมื่อล้มเหลววิธีกล่าวอ้างถึงความเสียเปรียบของตนเอง (Self – reported Handicaps) นำไปใช้โดย รายงานว่าพวกเขาได้รับอุปสรรคจากความเจ็บป่วย ความวิตกกังวล ความอาย เหตุการณ์ที่ทำให้เสียใจ เป็นต้น

แม้ว่ารูปแบบการทำให้ตนเองเสียเปรียบโดยการกระทำกับการกล่าวอ้างอุปสรรคของตนเองจะมีความแตกต่างกันแต่การศึกษาเหล่านี้ก็มีสมมติฐาน (assumption) ร่วมกันว่าคนที่เคยมีประสบการณ์ เกี่ยวกับความสำเร็จที่เกิดขึ้นในบางครั้งหรือมีความเกี่ยวข้องกับความล้มเหลวที่อาจจะเกิดขึ้นในอนาคตอาจเลือกกระทำในทางที่ไม่สามารถทำให้ผลงานนั้นออกมาได้ดีถ้าสภาพแวดล้อมสร้างความยากลำบาก มากกว่าจะกล่าวอ้างว่าไม่มีความสามารถ แต่ถ้าประสบความสำเร็จจะกล่าวอ้างว่ามาจากตนเอง

จอห์น และ โรดวอลด์ (Jones and Rhodewalt. 1982) ได้พัฒนามาตรวัดการทำให้ตนเองเสียเปรียบ (Self – handicapping Scale) เพื่อใช้วัดลักษณะนิสัยที่แต่ละคนมีแนวโน้มจะใช้วิธีทำให้ตนเองเสียเปรียบ เพื่อเป็นการป้องกันตนเอง บุคคลใดได้คะแนนสูงจากมาตรวัดนี้ มีแนวโน้มว่าจะมีพฤติกรรม ทำให้ตนเองเสียเปรียบมากกว่าบุคคลที่ทำคะแนนได้ต่ำ ซึ่งเห็นได้ว่าบุคลิกภาพการทำให้ตนเองเสียเปรียบสอดคล้องกับแนวคิดบุคลิกภาพของ โรเบิร์ต อี. ซิลเวอร์แมน (Robert E. Silverman) ที่ได้กล่าวถึง พฤติกรรมที่แสดงออกของบุคคล

ที่ไม่สามารถปรับตัวให้เข้ากับสิ่งแวดล้อมได้ คือ กลไกการป้องกันตัวเอง (defense mechanisms)

กลไกการป้องกันตนเอง (Defense Mechanism)

คนเราเกิดความคับข้องใจได้เสมอในชีวิตประจำวัน เช่น จะขับรถไปทำงานตอนเช้า แต่ติดเครื่องยนต์ไม่ได้ หรือเล่นการพนันแล้วเสีย หรือไม่ได้รับรางวัลตอบแทนจากผลการทำงานของตน เป็นต้น ในตอนเด็กๆ เราพบว่าความบกพร่องของเราเอง หรือไม่ก็สิ่งแวดล้อมที่ไม่เอื้ออำนวยเราจะขัดขวางการบรรลุเป้าหมายของเราเสมอ ปฏิกริยาของเราที่มีต่อความคับข้องใจต่างๆ จะช่วยให้เราปรับตัวต่ออุปสรรคเหล่านั้นได้ แต่เมื่อความคับข้องใจสะสมมากๆ เข้าและยากต่อการแก้ไขโดยตรง เราก็อาจมีปฏิกริยาเพื่อการป้องกันตนเองจากความรู้สึกที่เกิดขึ้นควบคู่กับความคับข้องใจ เราศึกษาวิธีที่ลดความไม่สบายใจที่เกิดขึ้นในลักษณะต่างๆ เสีย วิธีลดความเครียด ความไม่สบายใจ หรือเรียกได้ว่าเป็นปฏิกริยาตอบโต้ (reactions) เหล่านี้ เรียกว่ากลไกการป้องกันตนเอง (defense mechanism) ซึ่งจะถูกใช้เป็นเครื่องมือในการบิดเบือนหรือหนีความจริงที่สร้างความไม่สบายใจ หรือความทุกข์ให้แก่เรา

เราแยกชนิดของกลไกการป้องกันตนเอง ออกเป็น 2 ประเภท คือ ประเภทถอยหนี (escape techniques) และประเภทประนีประนอม (compromise techniques) ความก้าวร้าว (aggressive) นั้นในบางครั้งก็ถูกรวมเข้าไปเป็นกลไกการป้องกันตนเองชนิดหนึ่ง เพราะมันเป็นปฏิกริยาตอบโต้ชนิดหนึ่งต่อความคับข้องใจหรือความวิตกกังวล แต่อย่างใดก็ตีพฤติกรรมก้าวร้าวมักถูกใช้เป็นเครื่องตอบโต้โดยตรงต่อต้นตอของความคับข้องใจ (เช่น การขว้างปาสิ่งของไปยังบุคคลที่ตำหนิหรือทำร้ายเรา) ดังนั้น จึงมีลักษณะที่แตกต่างไปจาก การถอยหนีและการประนีประนอมดังกล่าวข้างต้น

1. กลไกการป้องกันตนเองประเภทถอยหนี (Escape techniques)

กลไกการป้องกันตนเองประเภทถอยหนีนี้มีหลายชนิด คือ การเก็บกด (repression) การสร้างจินตนาการ และการถอยกลับ (regression) กลไกประเภทนี้ใช้เพื่อหนีหรือหลีกเลี่ยง จากสถานการณ์ที่ทำให้เกิดความวิตกกังวล

การเก็บกด (repression) กลไกการป้องกันตัวที่ธรรมดา ๆ ทำได้ง่าย ก็คือการถอยหนีหลบไปจากสถานการณ์ที่ทำให้คับข้องใจหรือวิตกกังวลเสีย การเก็บกด ก็เป็นวิธีหนึ่งในหลายๆ วิธีของการถอยหนี Freud เป็นคนริเริ่มแนะนำคำๆ นี้ โดยอธิบายว่า หมายถึงกรรมวิธีของบุคคลที่พยายามฝังความคิดที่ ก่อให้เกิดความวิตกกังวลให้อยู่แต่ในจิตใจได้สนิท ไม่ให้ผุดขึ้นมาในจิตสำนึกและรบกวนชีวิตประจำวันได้ ตัวอย่างเช่น เมื่อครั้งก่อนเราไปหาหมอฟัน เราได้รับความเจ็บปวดทรมานมากและกลับเจ็บปวดที่บ้านอีกเป็นเวลานานหลายชั่วโมง พอถึงนัดคราวต่อไปซึ่งเป็นอาทิตย์หลังจากนั้นเราได้ "ลืม" ที่จะไปหาหมอดีก ตามนัด ปรากฏการณ์เช่นนี้เกิดขึ้นอาจเป็นเพราะว่า เราได้ "เก็บกด" การนึกถึงนัดเอาไว้เพื่อให้วิตกกังวลน้อยลง

จนเวลาผ่านไปโดยไม่รู้ตัว การเก็บกดที่เป็นลักษณะเด่นชัดและมีความรุนแรงที่สุดนั้นอยู่ในรูปของโรคลืม (amnesia) ซึ่งมีอาการหนักเบา หลายระดับขึ้นอยู่กับบุคคลและเหตุที่เกิดขึ้น ตัวอย่างเช่น การไปหาหมอฟันที่ดังที่กล่าวมาแล้ว บุคคลอาจลืมร้านว่าอยู่ที่ไหน ลืมเวลานัด และยิ่งกว่านั้นลืมชื่อหมอด้วย และหากความวิตกกังวลนั้นมีมาก คนๆ นั้นก็อาจไม่ลืมแต่ส่วนประกอบของเหตุการณ์เท่านั้นแต่ลืม ไปเลยว่า ตนเองได้มีความเจ็บป่วยนั้นๆ มาก่อน

การสร้างจินตนาการ (fantasy) บุคคลที่เกิดความคับข้องใจจากเหตุการณ์เป็นจริงที่ตนกำลังประสบอยู่ อาจหลบหนีความเป็นจริงเข้าสู่โลกแห่งจินตนาการ ซึ่งเป็นโลกที่ปราศจากความคับข้องใจที่จะมารบกวนเขา เขาไม่สามารถเผชิญหน้าและแก้ไขความทุกข์ได้ ก็เลยสร้างจินตนาการวาดมโนภาพขึ้นว่าเขาได้อยู่ในสภาวะที่เป็นสุขเป็นการลดความเครียดและสนองความต้องการของตนบุคคลที่ไม่สามารถ ปรับตัวให้เข้ากับความจริงในชีวิตประจำวันของตนได้ อาจหนีเข้าสู่โลกแห่งจินตนาการเลยทีเดียวและเมื่อหากเป็นเช่นนั้นไปแล้ว เขาก็จะต้องพึ่งพาอาศัยจินตนาการเป็นการแก้ปัญหาอยู่ร่ำไป จนในที่สุดไม่สามารถแยกความจริงกับจินตนาการออกจากกันได้ และเกิดพยาธิสภาพทางพฤติกรรมขึ้น

การถอยกลับ (regression) บุคคลอาจหนีจากความคับข้องใจ หรือสถานการณ์ที่ทำให้เกิด การวิตกกังวลได้อีกวิธีหนึ่ง คือ การถอยกลับไปใช้พฤติกรรมที่เคยใช้มาตอนเป็นเด็ก หรือพฤติกรรมแบบ คนโบราณ ตัวอย่างง่ายๆ ก็เช่นหญิงที่แต่งงานแล้วไม่สามารถจะปรับตัวให้เข้ากับบทบาทของภรรยาได้ เลยกลับบ้านไปหาพ่อแม่เพื่อที่จะได้กลับไปแสดงบทบาทของการเป็นลูกที่เคยอบอุ่นและเป็นสุข นักจิตวิทยาบางคนเชื่อว่า การแสดงตนว่าป่วยไข้ของบางคนเพื่อดึงดูดความสนใจของคนอื่น (hypochondriac) นั้น เป็นรูปหนึ่งของการถอยกลับ เพราะคนพวกนี้ใช้การป่วยไข้เป็นเครื่องดึงให้คนอื่น ๆ มาช่วยเหลือสนใจ ตนเหมือนกับเด็กๆ ที่ต้องพึ่งพาอาศัยพ่อแม่ การจะพึ่งพาอาศัยคนอื่นจะได้ง่ายขึ้นเมื่อเราแสดงตัวว่าเราป่วย การถอยกลับมักเกิดควบคู่กับการกระทำพฤติกรรมบางอย่างเป็นประจำ (stereotyped behavior) เช่น การดูดนิ้ว และทั้งสองอย่างทำให้บุคคลไม่สามารถจะปรับตัวให้เข้ากับสิ่งแวดล้อมของตนได้ พฤติกรรมที่ทำเป็นประจำ (stereotyped behavior) จะเกิดขึ้นเนื่องจากความคับข้องใจที่รุนแรง และการคงอยู่ของพฤติกรรมนั้น ทำให้บุคคลแก้ไขความคับข้องใจลำบากมากขึ้น

2. กลไกการป้องกันตนเองประเภทประนีประนอม (Compromise techniques)

กลไกประเภทนี้ ซึ่งมีการให้เหตุผล การชดโโทษ การทดแทนด้วยสิ่งที่สังคมยอมรับ การแปรเปลี่ยนปฏิกิริยา และการทดแทนด้วยสิ่งที่ตนสามารถทำได้เหล่านี้ ทำให้บุคคลสามารถจัดการกับสถานการณ์ที่ ทำให้เกิดความวิตกกังวลได้ วิธีการของเขา คือ เปลี่ยนสถานการณ์เสียบ้างให้ทิศทางเบนไปจากเดิมเท่าที่จะทำได้

การหาเหตุผลเข้าข้างตนเอง (rationalization) คนเราทุกคนมีความรักในคุณค่าของตน เราไม่อยากจะเป็นคนไร้ค่า ไม่ยอมเป็นที่ติฉินนินทาของสังคม แต่ในบางครั้ง เราก็ได้ทำอะไรที่น่าตำหนิลงไปแล้วด้วยความจงใจ หรือไม่จงใจก็ตาม เราก็เกิดความรู้สึกไม่สบายใจ และพยายามปกป้องตัวเองโดยการหาเหตุผลมาแก้ตัว หรือกล่าวหาคนอื่นและบางครั้งก็กล่าวหาสถานการณ์ว่าเป็นต้นเหตุของความผิดพลาด แทนที่จะยอมรับว่าตนเป็นต้นเหตุ ลักษณะการแก้ตัวเช่นนี้ เราเรียกว่าการให้เหตุผล (rationalization) โดยเหตุที่คนเรามักไม่อยากจะยอมรับความล้มเหลวของตัวเอง เมื่อใดที่ประสบความล้มเหลวเราก็พยายามที่จะกลบเกลื่อนโดยการบอกว่าที่จริงแล้วเราไม่ได้ล้มเหลวหากแต่เราเกิดไม่ต้องการทำงานนั้นต่อ หรือสิ่งที่เรากำลังทำนั้นไม่ดี ไม่คุ้มค่าพอที่จะพยายามต่างหากเมื่อเราทำไม่ดีต่อคนอื่น เรากลัวถูกสังคมตำหนิ เราก็เลยพยายามให้เหตุผลว่า คนที่ถูกเรากระทำนั้นได้ดี สมควรถูกลงโทษ เมื่อเราได้คะแนนในการสอบไม่ดี เราก็บอกว่าพวกคนอื่นที่สอบได้ดี เป็นพวกบ้าหนังสือไม่รู้จักใช้เวลาให้สนุกสนาน เมื่อเราสมัครงานแล้วไม่ได้ เราก็บอกว่าอันที่จริงแล้วเราไม่อาจได้งานนั้นเท่าไรหรอก ฯลฯ ตามหลักแล้ว การให้เหตุผลเหล่านี้เป็น วิธีการปกปิดในการที่จะได้ความคับข้องใจของคนทั่วไป เราใช้มันเพื่อลดความเครียดอันเกิดจากความ ผิดพลาดหรือความล้มเหลวของเราเอง แต่การใช้ปริมาณที่มากและบ่อยครั้งเกินไป ก็เป็นอาการที่บ่งบอกถึงความผิดปกติของพฤติกรรมที่รุนแรงได้

การชัดโทษ (Projection) บุคคลที่ป้ายความบกพร่องของตนให้กับผู้อื่นนั้น เป็นคนที่กำลังใช้กลไกป้องกันตนเองแบบที่เรียกว่า การชัดโทษ (projection) การเก็บกดและปิดบังลักษณะที่ไม่เหมาะสมของตัวเองและขณะเดียวกันก็ป้ายลักษณะที่ไม่เหมาะสมเหล่านั้นไปอยู่ในคนอื่นจะทำให้บุคคลรู้สึกสบายใจขึ้น ลักษณะที่ไม่เหมาะสมเหล่านั้นก็มีตัวอย่าง ความรู้สึกที่ไม่ดีต่าง ๆ ที่คนมักมีอยู่ในใจความขี้อิจฉาริษยา ความก้าวร้าวมุ่งร้าย หรือความรักที่น่าอับอาย เป็นต้น

การทดแทน (sublimation) ได้เคยกล่าวมาแล้ว กลไกการป้องกันตัวเองบางอันไม่ได้ใช้สำหรับปิดกั้นหรือหลีกเลี่ยงสถานการณ์ที่คับข้องใจ แต่ใช้สำหรับแก้ความคับข้องใจทางอ้อม Freud เคยพูดถึง การทดแทน (sublimation) เอาไว้ว่า เป็นการที่คนสร้างจุดมุ่งหมายที่สองขึ้นมาทดแทนจุดมุ่งหมายที่พลาด เมื่อจุดมุ่งหมายแรก บุคคลนั้นต้องการเกิดบรรลุไม่ได้ จะเนื่องด้วยจุดมุ่งหมายที่สังคมไม่ยอมรับหรือเป็นจุดมุ่งหมายที่ตนไม่สามารถที่จะเอามาได้ก็ตาม เขาก็เลยเปลี่ยนทิศทางของพฤติกรรมให้มุ่งไปยัง จุดหมายใหม่ที่เขาสามารถจะไปถึงได้โดยไม่รู้สึกว้าเหวหรือเสียหาย

การกลบเกลื่อนโดยแสดงออกในทางตรงกันข้าม (reaction formation) กลไกการป้องกันตัวอีกแบบหนึ่งซึ่งมีลักษณะคล้ายกันมากกับการทดแทน ก็คือ การกลบเกลื่อนโดยทำสิ่งที่ตรงข้ามกับที่อยากจะทำ เกิดขึ้นเมื่อบุคคลพยายามป้องกันตัวเองจากความรู้สึกที่เก็บกดเอาไว้โดยการสร้างความเชื่อและการกระทำที่ตรงกันข้ามกับความรู้สึกที่มีอยู่จริง

การกลบเกลื่อนโดยแสดงออกในทางตรงกันข้ามแบบนี้มักฝังแน่นอยู่ลึกในจิตใต้สำนึก และยากต่อการค้นพบ ในบางโอกาสมันอาจก่อให้เกิดประโยชน์มากกว่าโทษ อย่างเช่น ในกรณีที่ถูกใช้เพื่อป้องกัน พฤติกรรมที่ไม่ดี แต่ในบางกรณีมันก็อาจก่อให้เกิดโทษ ทั้งแต่ตนเองและผู้อื่นได้

การเบนเป้าหมาย (compensation) เมื่อคนเราเกิดความคับข้องใจอันเนื่องมาจากความล้มเหลวหรือเสียคุณค่าในตัวเนื่องจากการกระทำที่ผิดพลาด เราก็มักจะพยายามหาเป้าหมายใหม่ที่คิดว่าจะไม่ทำให้ตนล้มเหลวผิดพลาด โดยส่วนใหญ่แล้วเป้าหมายใหม่นี้จะมีลักษณะใกล้เคียงกับเป้าหมายเดิมของตัว เช่น คนที่ไฝฝืนอยากเป็นนักเขียนที่ยิ่งใหญ่ เมื่อเป็นไม่ได้ก็เลยไปประกอบอาชีพเป็นนักวาดคำโฆษณา การเบนเป้าหมาย (compensation) มีลักษณะคล้ายการทดแทน (sublimation) คือ ทำให้บุคคลสามารถทดแทนเป้าหมายหนึ่งด้วยอีกเป้าหมายหนึ่งได้ แตกต่างกันตรงที่การทดแทน (sublimation) เกิดจากความคิดหรือความรู้สึกว่าไม่เป็นที่ยอมรับของสังคม ส่วนการเบนเป้าหมาย (compensation) เกิดจากความล้มเหลวที่เกิดขึ้นแล้ว

อย่างไรก็ตาม ความพยายามที่จะเบนเป้าหมายมากจนเกินไป (over compensation) ทำให้เกิดผลเสียหายนามากกว่าที่จะเกิดผลดี เช่น คนที่รู้สึกว่าคุณค่าตนมีปมด้อยมากเหลือเกิน ความรู้สึกด้อยดำนี้อาจมีเสียดแทงใจเขาเลยพยายามที่จะหาการยอมรับจากคนอื่นมากขึ้นด้วยการแสดงความเด่นของตนมากจนเกินไป การเบนเป้าหมายมากจนเกินไปเป็นที่มาของความรำคาญ ของผู้อื่น และทำให้ผู้อื่นไม่ยอมรับ คนที่มีอาการเบนเป้าหมายมากจนเกินไป (over compensation) อาจได้รับสมญาต่อไปว่าเป็นคน “ขี้อ้อ” “ปากมาก” หรือ “นักรำคาญ” ได้ (บัวทอง สว่างโสภาคกุล. 2529 : 282 – 290)

จากที่กล่าวมา สามารถสรุปได้ว่าทั้งกลไกการป้องกันตนเอง (Defense Mechanism) และการทำให้ตนเองเสียเปรียบ (Self handicapping) เป็นวิธีการทำให้บุคคลอื่นเข้าใจว่าความล้มเหลวที่เกิดขึ้นนั้นเกิดจากสภาวะแวดล้อมต่างๆ ขัดขวางการบรรลุความสำเร็จของตนเอง ซึ่งถือเป็นวิธีปกปิดในการใช้เป็นเครื่องมือทางจิตที่เราใช้โต้ตอบความคับข้องใจ ความวิตกกังวล ความเครียด และความไม่สบายใจต่างๆ ขอบคนทั่วไปที่นับถือตัวเองต่ำ หรือผู้ที่ไม่แน่ใจระดับความสามารถที่แท้จริงของตนเอง

3. ความเชื่อมั่น (Reliability)

3.1 ความหมายของความเชื่อมั่น

นักการศึกษาและนักจิตวิทยาได้ให้ความหมายของความเชื่อมั่นไว้ต่างกัน ดังนี้
 นันนอลลี (Nunnally, 1964 : 59) กล่าวว่า ความเชื่อมั่น คือ สัดส่วนระหว่าง
 ความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการทดสอบ

ลอร์ดและโนวิก (Lord and Novick, 1968 : 198) กล่าวว่า ความเชื่อมั่น
 คือ ความคงที่ของคะแนนที่ได้จากการตอบแบบทดสอบซ้ำ และคะแนนที่ได้จากการตอบแบบ
 ทดสอบทั้งสองครั้ง เป็นอิสระไม่ขึ้นกับความคลาดเคลื่อนของการวัดใดๆ

ลินด์วอลล์ และนิคโค (Lindvall and Nitko, 1967 : 126) กล่าวว่า ความเชื่อมั่น
 คือค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบสองครั้งโดยใช้แบบทดสอบฉบับเดียวกันสอบ
 ในเวลาที่ต่างกัน

อนาสตาซี (Anastasi, 1968 : 105) กล่าวว่า ความเชื่อมั่น หมายถึง ความคงที่
 แน่นนอนของคะแนนซึ่งได้จากการวัดนักเรียนกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวกัน
 หลายๆ ครั้ง หรือด้วยแบบสอบสองฉบับที่มีลักษณะเสมอเหมือนกัน หรือภายใต้เงื่อนไขของ
 ตัวแปรอื่นๆ ในการวัดนั้น

วิรัช วรรณรัตน์ (2532 : 53) กล่าวว่า ความเชื่อมั่น หมายถึง ความสามารถของ
 เครื่องมือที่สามารถให้ผลการวัดที่ถูกต้องแน่นอน (Accuracy) คงเส้นคงวา (Consistency) เป็น
 ที่เชื่อถือ หรือมั่นใจในผลที่ได้ออกมาอย่างแท้จริง แม้จะมีการวัดซ้ำอีก ผลที่ได้ย่อมไม่เปลี่ยนแปลง
 ไปจากเดิม

บุญเชิด ภิญโญนนตพงษ์ (2545 : 198) กล่าวว่า ความเชื่อมั่น หมายถึง การวัด
 มีความอิสระหรือปลอดจากความคลาดเคลื่อนได้ผลการวัดที่สอดคล้องกันซึ่งเป็นการแสดงถึง
 ขนาดของความแปรปรวนจริงและความแปรปรวนคลาดเคลื่อน หรือกล่าวได้ว่าเป็นขนาดอิทธิพล
 ของขนาดองค์ประกอบความคลาดเคลื่อนแบบสุ่มหรือไม่เป็นระบบ ของการวัดที่จะทำให้ผลการวัด
 มีความสอดคล้องกันมากน้อยเพียงไร

จากที่กล่าวมา สรุปได้ว่า ความเชื่อมั่น หมายถึง ความคงที่แน่นอนของคะแนน
 ที่ได้จากการสอบวัดจากคนกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวหรือขนานกัน ในเวลาที่ต่างกัน
 หรืออีกนัยหนึ่งก็คือสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของ
 คะแนนที่ได้จากการทดสอบ

3.2 ทฤษฎีความเชื่อมั่น (Theory of Reliability)

ทฤษฎีความเชื่อมั่น กล่าวว่า คะแนนที่ได้จากการทดสอบ (observed Score) จะประกอบด้วยคะแนนจริง (True Score) และคะแนนความคลาดเคลื่อน (Error Score) ดังสมการ

$$X = T + E \quad \dots\dots\dots (1)$$

เมื่อ X = คะแนนที่ได้จากการทดสอบ

T = คะแนนจริง

E = คะแนนความคลาดเคลื่อน

คะแนนจริง (True Score) หมายถึง คะแนนที่ผู้สอบได้รับจากการวัดด้วยเครื่องมือที่มีคุณภาพสูง ปราศจากความคลาดเคลื่อน หรือ หมายถึง คะแนนเฉลี่ยของผู้สอบซึ่งได้จากการทำแบบทดสอบฉบับเดิมหลายๆ ครั้ง โดยมีข้อตกลงว่าจะต้องไม่มีอิทธิพลจากการฝึกฝนความเมื่อยล้าและการเรียนรู้ในการทดสอบซ้ำ

คะแนนความคลาดเคลื่อนคือ คะแนนที่เกิดจากความคลาดเคลื่อนในการวัดแบ่งออกเป็น 2 ประเภท คือ ความคลาดเคลื่อนที่เกิดจากการสุ่ม (Random Error) เป็นความคลาดเคลื่อนที่เกิดขึ้น โดยบังเอิญกับความคลาดเคลื่อนอย่างมีระบบ (Systematic Error) สำหรับความคลาดเคลื่อนประเภทหลังจะ ไม่มีผลกระทบต่อความเชื่อมั่น ส่วนความคลาดเคลื่อนที่เกิดขึ้นโดยการสุ่มอาจเป็นไปได้ทั้งทางบวกและ ทางลบ เมื่อรวมกันแล้วมีค่าเป็นศูนย์ ($\sum E = 0$) และถ้าความคลาดเคลื่อนที่เกิดขึ้นเป็นไปในลักษณะสุ่มแล้ว ก็จะไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือ สหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนมีค่าเท่ากับ 0 ($r_{TE} = 0$) ดังนั้น ความแปรปรวนของคะแนนที่วัดได้สามารถเขียนได้ดังสมการ (Mehrens and Lehmann. 1973 : 91) ดังนี้

$$S_x^2 = S_T^2 + S_E^2 \quad \dots\dots\dots (2)$$

เมื่อ S_x^2 แทน ค่าความแปรปรวนของคะแนนที่ได้จากการสอบ

S_T^2 แทน ค่าความแปรปรวนของคะแนนจริง

S_E^2 แทน ค่าความแปรปรวนของคะแนนความคลาดเคลื่อน

จากทฤษฎีความเชื่อมั่น (r_{tt}) คืออัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สอบ

$$\text{นั่นคือ } r_{tt} = \frac{S_T^2}{S_X^2} \dots\dots\dots (3)$$

จากสมการ (2) และ (3) จะได้สมการพื้นฐานในการหาสัมประสิทธิ์สหสัมพันธ์ของความเชื่อมั่น ดังนี้

$$r_{tt} = 1 - \frac{S_E^2}{S_X^2}$$

(Mehrens and Lehmann. 1984 : 268 – 269)

เมื่อ	r_{tt}	แทน	ความเชื่อมั่นของแบบทดสอบ
	S_T^2	แทน	ค่าความแปรปรวนของคะแนนจริงที่ได้จากการสอบ
	S_E^2	แทน	ค่าความแปรปรวนของคะแนนความคลาดเคลื่อน
	S_X^2	แทน	ค่าความแปรปรวนของคะแนนที่ได้จากการสอบ

3.3 วิธีการประมาณค่าความเชื่อมั่น

เฟอร์กูสัน (Ferguson. 1966 : 365 – 366) และสแตนเลย์ และฮอปกินส์ (Standley and Hopkins. 1972 : 122 – 127) ได้กล่าวในลักษณะเดียวกันว่ามีวิธีการหาสัมประสิทธิ์ความเชื่อมั่น 4 วิธี ดังนี้

1. วิธีสอบซ้ำ (Test Retest Method) หรือบางครั้งเรียกว่า สัมประสิทธิ์ของความคงที่ (Coefficient of Stability) เป็นการนำแบบทดสอบฉบับเดียว ไปทำการทดสอบกับบุคคลเดียวกันซ้ำสองครั้งในช่วงเวลาที่แตกต่างกันพอสมควร คะแนนที่ได้จากการทำแบบทดสอบทั้งสองครั้งมีสหสัมพันธ์กัน ค่าสหสัมพันธ์ที่ได้เป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

2. วิธีใช้แบบทดสอบคู่ขนาน (Parallel Forms Method) เป็นการนำแบบทดสอบที่มีลักษณะ คู่ขนานกันหรือเท่าเทียมกัน โดยมีเนื้อหา ค่าเฉลี่ย และค่าความแปรปรวนเท่ากันไปทดสอบในเวลาเดียวกันหรือเวลาที่แตกต่างกัน คะแนนที่ได้จากแบบทดสอบทั้งสองฉบับมีสหสัมพันธ์ ค่าสหสัมพันธ์ที่ได้เป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

3. วิธีแบ่งครึ่งแบบทดสอบ (Split-Half Method) เป็นการนำแบบทดสอบฉบับเดียว ไปทดสอบกับบุคคลกลุ่มเดียว แล้วแบ่งครึ่งแบบทดสอบเป็นชุดของคะแนนข้อคู่ และชุดคะแนนข้อคี่แล้วนำคะแนนที่ได้จากการแบ่งครึ่งแบบทดสอบไปหาสหสัมพันธ์กัน จากนั้นปรับขยายด้วยสูตรของสเปียร์แมน-บราวน์ เป็นสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4. วิธีวัดความคงที่ภายในของแบบทดสอบ (Internal-Consistency Method) เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มเดียว แล้วนำไปหาสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบโดยวิธีของคูเดอร์ – ริชาร์ดสัน (Kuder-Richardson)

อนาสตาเซีย (Anastasi. 1976 : 79-86) ยังได้เสนอวิธีการหาค่าความเชื่อมั่นไว้ดังนี้

1. ความเชื่อมั่นแบบสอบซ้ำ (Test-Retest Reliability) เป็นวิธีที่ชัดเจนที่สุดที่จะหาว่า แบบทดสอบนั้นเชื่อถือได้หรือไม่ ก็คือใช้การทดสอบหาค่าความเชื่อมั่นแบบสอบซ้ำ สัมประสิทธิ์แห่งความเชื่อมั่นในกรณีนี้คือสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสองครั้งที่ได้จากการทดสอบคนกลุ่มเดียวกัน ด้วยแบบทดสอบเดิมสองครั้ง ความแปรปรวนคลาดเคลื่อนจะเป็นการเปลี่ยนแปลงที่ไม่ได้ตั้งใจให้เกิดขึ้นเนื่องจากเวลาในการทดสอบต่างกัน ในบางกรณีความแปรปรวนดังกล่าวอาจจะเนื่องมาจากสภาพที่ควบคุมไม่ได้ เช่น ผู้รับการทดสอบเกิดป่วย เหนื่อย อารมณ์เครียด กังวล ตีใจหรือไม่สบายใจ ความเปลี่ยนแปลงของอากาศ เสียงและเครื่องล่อใจอื่นๆ หรือดินสอหัก เป็นต้น ความเชื่อมั่นแบบสอบซ้ำมีค่าสูง แสดงว่าสิ่งแวดล้อมที่เปลี่ยนแปลงเป็นประจำนั้นมีผลน้อยต่อคะแนนสอบ ซึ่งแบบทดสอบจิตวิทยาส่วนใหญ่ไม่เหมาะสมที่จะใช้เทคนิคการสอบซ้ำในการหาความเชื่อมั่นของแบบทดสอบดังกล่าว เนื่องจากผลการตอบข้อสอบครั้งแรกอาจจะทำให้ผู้รับการทดสอบตอบข้อสอบครั้งที่สองได้ดีขึ้นตามความแตกต่างของคน ยิ่งกว่านั้นถ้าระยะของการทดสอบใกล้กันมาก ผู้ตอบก็อาจจะยังคงจำแบบของข้อสอบได้ เมื่อได้รับการทดสอบซ้ำ ดังนั้นคะแนนที่ได้รับจากการทดสอบทั้งสองครั้งจะเกี่ยวข้องกันมาก จนทำให้ค่าสัมประสิทธิ์สหสัมพันธ์สูงอย่างไม่น่าเชื่อ ยิ่งถ้าเป็นการทดสอบเกี่ยวกับปัญหาที่ต้องใช้เหตุผลหรือความคิดใหม่ๆ แล้วเขาก็จะตอบปัญหาเดิมนั้นได้โดยไม่ต้องผ่านการคิดบางขั้นเลย ในการทดสอบครั้งที่สอง แบบทดสอบซึ่งปลอดจากอิทธิพลของการทำซ้ำเท่านั้นจึงจะใช้เทคนิคการสอบซ้ำได้ดี

2. ความเชื่อมั่นแบบใช้สอบคนละฉบับ (Alternate – Form Reliability) เนื่องจากการหาความเชื่อมั่นแบบสอบซ้ำมีความยุ่งยากคือต้องทำการสอบถึง 2 ครั้ง จึงทำการหลีกเลี่ยงความยุ่งยากโดยใช้แบบทดสอบสองฟอร์มแทนแล้วหาค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนจากการทดสอบด้วยแบบทดสอบดังกล่าวสองฟอร์ม สัมประสิทธิ์สหสัมพันธ์นี้จะเป็นสัมประสิทธิ์แห่งความเชื่อมั่น ซึ่งเป็นทั้งความคงที่ทางเวลา และความสอดคล้องกันของการตอบสนองต่อข้อคำถามหรือแบบทดสอบที่ต่างกัน ดังนั้นสัมประสิทธิ์อย่างนี้จึงประกอบด้วยความเชื่อมั่นสองแบบรวมกัน ความเชื่อมั่นแบบใช้สอบคนละฉบับให้ประโยชน์แก่การวัดเพื่อประเมินค่าของแบบทดสอบชนิดต่างๆ ในการสร้างแบบทดสอบใช้สอบคนละฉบับต้องแน่ใจว่าเป็นฟอร์ม คู่ขนานกันจริงๆ หลักเบื้องต้นก็คือสร้างฟอร์มคู่ขนานของแบบทดสอบโดยอาศัยตารางกำหนดงานเดียวกัน แบบทดสอบควรมีจำนวนข้อเท่ากันวางรูปแบบเหมือนกันและครอบคลุมเนื้อหาอย่างเดียวกัน พิสัย

และระดับของความยากควรเท่ากัน คำสั่งจำกัดเวลา ตัวอย่างทุกอย่างต้องเหมือนกันและเทียบกันได้

แม้ว่าความเชื่อมั่นแบบใช้สอบคนละฉบับจะใช้ได้มากกว่าแบบสอบซ้ำ ก็ยังมีข้อบกพร่องบางประการ ประการแรกก็คือ ถ้าพิจารณาถึงเหตุการณ์ที่เกิดขึ้นเนื่องจากการฝึกแบบทดสอบใช้สอบคนละฉบับจะลดผลการเปลี่ยนแปลงเนื่องจากการฝึก แต่จะไม่จำกัดผลการเปลี่ยนแปลงดังกล่าวได้ เพื่อความแน่ใจจึงอธิบายต่อไปได้ว่า ถ้าผู้รับการทดสอบทุกคนมีความสามารถเพิ่มขึ้นเท่ากันเมื่อทดสอบฟอร์มที่สอง ความสัมพันธ์ของคะแนนจากการทดสอบใช้สอบคนละฉบับก็จะคงเดิม เพราะเท่ากับเพิ่มคะแนนจำนวนคงที่ให้แก่ทุกคนซึ่งไม่ทำให้ ค่าสัมประสิทธิ์สหสัมพันธ์เปลี่ยนแปลงไป อย่างไรก็ตามก็ตีบุคคลจะพัฒนามากน้อยต่างกัน เนื่องจากขอบเขตแห่งการฝึกมาก่อนด้วยวัสดุที่คล้ายกัน สิ่งเราในการทดสอบและตัวประกอบอื่นๆ ด้วย โดยสภาพดังกล่าวเป็นผลมาจากการฝึกจะเป็นแหล่งแห่งความแปรปรวน ซึ่งแนวโน้มที่จะลดค่าสหสัมพันธ์ระหว่างการทดสอบด้วย แบบทดสอบใช้สอบคนละฉบับ แต่ถ้าปัญหาเกี่ยวกับฝึกมีน้อย ค่าสหสัมพันธ์ก็แทบจะไม่ลดน้อยเลย

3. ความเชื่อมั่นชนิดแบ่งครึ่ง (Split – Half Reliability) เป็นการวัดความเชื่อมั่นของแบบทดสอบโดยการทดสอบเพียงครั้งเดียวด้วยแบบทดสอบฟอร์มเดียว แล้วแบ่งข้อสอบออกเป็น 2 ส่วน ด้วยวิธีการต่างๆ โดยวิธีนี้ผู้รับการทดสอบทุกคนจะได้รับคะแนนสองจำนวนจากการแบ่งแบบทดสอบเป็นสองส่วนที่พอเทียบกันได้ เห็นได้ชัดว่าความเชื่อมั่นแบบนี้วัดความสม่ำเสมอที่เกี่ยวกับเนื้อหาที่สุ่มมา ความคงที่ของคะแนนเนื่องจากเวลาจะไม่มาเกี่ยวกับความเชื่อมั่นแบบนี้ เพราะการทำการทดสอบครั้งเดียว สัมประสิทธิ์แห่ง ความเชื่อมั่นแบบนี้บางทีเรียกว่า สัมประสิทธิ์แห่งความสอดคล้องภายใน เพราะดำเนินการสอบครั้งเดียวด้วยแบบทดสอบฉบับเดียว

ในการหาความเชื่อมั่นแบบนี้ปัญหาประการแรกที่พบ คือจะแบ่งแบบทดสอบอย่างไรจึงจะได้ ข้อสอบที่ทั้งสองส่วนเทียบกันได้มากที่สุด อาจจะปฏิบัติได้หลายทาง ข้อสอบครั้งแรกและครึ่งหลังส่วนมากจะเทียบกันไม่ได้เนื่องจากความยากไม่ได้ขนาดเดียวกัน ผลจากการอ่านเครื่อง การฝึกความล้า ความเบื่อ และตัวประกอบอื่นๆ ที่แปรไปตามความก้าวหน้าจากการเริ่มต้นต่อไปจนจบ ส่วนใหญ่จะคิดคะแนนรวมของข้อใด ส่วนหนึ่งและข้อคู่อีกส่วนหนึ่ง ถ้าโดยประมาณแล้ว ข้อสอบเริ่มต้นเรื่องความยากตามระเบียบ การแบ่งอย่างนี้ก็จะได้จากแต่ละครั้งแบบทดสอบที่เกือบจะสมมูลย์กัน ข้อควรสังเกตก็คือ การแบ่งข้อดีและข้อคู่ของกลุ่ม ข้อสอบที่เป็นปัญหาอย่างเดียวกัน เช่น เรื่องที่ต้องแก้ไขไดอาแกรมทางเมคานิกส์หรือข้อความที่กำหนดให้ในการทดสอบการอ่าน ในกรณีนี้ข้อสอบทั้งหมดควรมีคำถามครบถ้วนในทั้งสองครึ่ง เมื่อแบ่งข้อสอบในกลุ่มเช่นนั้นออกเป็นสองส่วน คะแนนจากทั้งสองส่วนก็จะเพิ่มความคล้ายกันขึ้น เพราะเข้าใจว่าแบบทดสอบทั้งสองส่วนจะมีความคลาดเคลื่อนอย่างเดียวกัน

4. ความเชื่อมั่นแบบคูเดอร์-ริชาร์ดสัน (Kuder-Richardson Reliability) สามารถหาได้โดยดำเนินการสอบครั้งเดียวและใช้แบบทดสอบชุดเดียว ซึ่งขึ้นอยู่กับความสม่ำเสมอของผู้รับการทดสอบต่อ ข้อสอบทุกข้อในแบบทดสอบ ความสอดคล้องกันระหว่างข้อนี้ได้รับอิทธิพลจากแหล่งแห่งความแปรปรวนคลาดเคลื่อน 2 ชนิด คือ เนื้อหาที่สุมมาและความเป็นวิวิธพันธ์ของพฤติกรรมที่สุมมา ยิ่งข้อสอบมีความเป็นเอกพันธ์มากขึ้นความสอดคล้องกันระหว่างข้อก็ยิ่งสูงขึ้น ในบรรดาสูตรที่คูเดอร์และริชาร์ดสันเขียนไว้สูตรที่ใช้มากที่สุดคือสูตร คูเดอร์-ริชาร์ดสัน 20

5. ความเชื่อมั่นของผู้ให้คะแนน (Scorer Reliability) หาได้โดยให้คน 2 คน ตรวจคำตอบใน ข้อสอบฉบับเดียวกันของผู้รับการสอบแต่ละคน โดยไม่เกี่ยวข้องกัน จะได้คะแนนสองชุดที่แสดงความสามารถของคนกลุ่มเดียวกัน วัดด้วยเครื่องมืออย่างเดียวกันแต่ตัดสินโดยคนสองคนโดยไม่เกี่ยวข้องกัน จากนั้นนำความสัมพันธ์ทั้งสองชุดมาหาความสัมพันธ์แบบธรรมดา

อัลเลนและเยน (Allen and Yen. 1979 : 88) ได้กล่าวว่า วิธีการประมาณค่าความเชื่อมั่นที่แตกต่างกันจะให้ความเชื่อมั่นที่ต่างกัน ในการพิจารณาค่าความเชื่อมั่นสำหรับแบบทดสอบที่อาศัยความเร็วนั้น ควรใช้แบบสอบซ้ำหรือแบบทดสอบคู่ขนาน เพราะว่าการวัดความสอดคล้องภายในส่วนใหญ่อาจจะทำให้เพื่อ ส่วนการใช้วิธีสัมประสิทธิ์แอลฟา (α) และวิธีของคูเดอร์ ริชาร์ดสัน ให้ผลการประมาณค่า ความเชื่อมั่นที่ต่ำ และใช้กับแบบทดสอบที่มีลักษณะเป็นเอกพันธ์ (Homogeneous) เท่านั้น เพราะว่าสูตรเหล่านี้มีพื้นฐานมาจากความเป็นเอกพันธ์ของข้อสอบ ถ้าเป็นข้อสอบที่วัดคุณลักษณะที่แตกต่างกันการหาความเชื่อมั่นแบบสัมประสิทธิ์สหสัมพันธ์ และคูเดอร์ ริชาร์ดสัน จะไม่เหมาะสมเพราะได้ค่าต่ำกว่าที่ควร

พวงรัตน์ ทวีรัตน์ (2543 : 122 – 127) ได้กล่าวถึงการตรวจสอบหรือหาค่าความเชื่อมั่นของวิธีการอยู่หลายวิธี แต่ละวิธีก็เหมาะสมกับเครื่องมือ แต่ละชนิด ทั้งนี้ขึ้นอยู่กับลักษณะของเครื่องมือและคุณลักษณะของสิ่งที่ต้องการศึกษา การหาค่าความเชื่อมั่นแต่ละวิธีมีดังต่อไปนี้

ก. แบบสอบซ้ำ (Test – retest method) เป็นการนำเครื่องมือที่สร้างขึ้น ไปสอบวัดกับคนกลุ่มเดียวกันสองครั้งในเวลาต่างกัน ได้คะแนนมาสองชุด นำคะแนนทั้งสองชุด มาหาค่าสัมประสิทธิ์สหสัมพันธ์ ถ้าเป็นคะแนนดิบก็ใช้วิธีของเพียร์สัน ถ้าเป็นคะแนนในรูปอื่นก็หาค่าสัมประสิทธิ์สหสัมพันธ์ด้วยวิธีอื่น ถ้าได้ค่าสัมประสิทธิ์สหสัมพันธ์สูงก็แสดงว่าเครื่องมือที่มีความเชื่อมั่นสูง การหาค่าความเชื่อมั่นแบบนี้เป็นการวัดความคงที่ภายนอก (Stability) สัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้อาจเรียกว่า สัมประสิทธิ์ของความคงที่ (Coefficient of stability)

การหาความเชื่อมั่นโดยวิธีนี้ มีข้อควรคำนึงถึง คือ ระยะเวลาที่ทั้งช่วงระหว่างการสอบสองครั้ง ไม่ควรห่างกันมากนัก หรือใกล้กันเกินไป เพราะจะมีผลกระทบต่อค่าความเชื่อมั่น กล่าวคือ ถ้าทั้งช่วงนานเกินไปกลุ่มตัวอย่างอาจเกิดความงอกงามทางการเรียนรู้ ทำให้สามารถทำคะแนนในครั้งหลังได้สูงขึ้น หรือถ้าระยะเวลาใกล้กันเกินไปกลุ่มตัวอย่างอาจจำข้อสอบได้ทำให้คะแนนได้สูงขึ้น หรือกลุ่มตัวอย่างอาจขาดความสนใจในการทำครั้งที่สองก็จะทำให้ค่าความเชื่อมั่นคลาดเคลื่อนไปได้เช่นกัน โดยทั่วไประยะที่ทั้งห่างกันของการสอบสองครั้งไม่ควรเกิน 3 สัปดาห์ แต่ทั้งนี้ขึ้นอยู่กับระดับอายุของกลุ่มตัวอย่างด้วย

ข. แบบใช้เครื่องมือวัดที่มีลักษณะเท่าเทียมกันหรือคู่ขนานกัน (Equivalent form or parallel form method) เป็นการหาความเชื่อมั่นโดยนำเครื่องมือที่สร้างขึ้นกับเครื่องมืออีกฉบับหนึ่งที่มีคุณภาพเหมือนกันทุกประการ คือ มีเนื้อหา รูปแบบข้อคำถาม จำนวนข้อความยากง่าย เหมือนกัน และมีค่าเฉลี่ยและความแปรปรวนเท่ากันทั้งสองฉบับไปสอบวัดกับกลุ่มทดลองเครื่องมือกลุ่มเดียวกัน ได้คะแนน 2 ชุด นำมาหาค่าสัมประสิทธิ์สหสัมพันธ์ตามวิธีของเพียร์สัน (ถ้าเป็นคะแนนดิบ) ค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้เรียกสัมประสิทธิ์ของความเท่าเทียมกัน (Coefficient of equivalence)

ค. แบบแบ่งครึ่ง (Split half method) เป็นการนำเครื่องมือที่ต้องการหาความเชื่อมั่น ไปสอบวัดกับกลุ่มทดลองเครื่องมือกลุ่มเดียว แล้วนำเครื่องมือชิ้นพร้อมคำตอบมาแบ่งครึ่งเป็น 2 ฉบับ ส่วนมากการแบ่งครึ่งมักจะแบ่งเป็นข้อคู่ ข้อคี่ ซึ่งเมื่อแบ่งแล้วจะได้ข้อสอบ 2 ฉบับ ซึ่งมีจำนวนข้อเท่ากัน ตรวจให้คะแนนข้อคู่ครั้งหนึ่ง และข้อคี่อีกครั้งหนึ่ง ได้คะแนน 2 ชุด สมาชิกในกลุ่มแต่ละกลุ่มแต่ละคนจะได้คะแนน 2 ตัว นำคะแนนทั้ง 2 ชุด นั้นมาหาค่าสัมประสิทธิ์สหสัมพันธ์ตามวิธีของเพียร์สัน ค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้เป็นค่าความเชื่อมั่นของเครื่องมือชิ้นเพียงครึ่งฉบับ ซึ่งต้องปรับให้เป็นค่าความ เชื่อมั่นทั้งฉบับ โดยใช้สูตรของสเปียร์แมน บราวน์ (Spearman Brown)

การหาความเชื่อมั่นของเครื่องมือโดยใช้วิธีแบบครึ่งนี้ เป็นการหาความคงที่หรือความคงตัวภายใน (Internal consistency)

ข้อสังเกต การหาความเชื่อมั่นของเครื่องมือแบบครึ่งนี้ เมื่อแบ่งออกเป็นสองส่วนแล้ว เนื้อหาของทั้งสองส่วนนั้นจะต้องมีลักษณะคู่ขนานกัน คือ ต้องคู่ขนานกันทั้งในสิ่งที่ถาม รูปแบบของคำถาม และความยากง่ายของข้อคำถาม ซึ่งมีลักษณะคล้ายข้อสอบคู่ขนานกันนั่นเอง ถ้าทำได้เช่นนี้จึงจะทำให้ค่าความเชื่อมั่นที่ได้มีความถูกต้องยิ่งขึ้น

ง. แบบของคูเดอร์ – ริชาร์ดสัน (Kuder – Richardson) การหาความเชื่อมั่นวิธีนี้เป็นที่นิยมมาก เพราะมีข้อดีตรงที่วัดสอบครั้งเดียวกับกลุ่มตัวอย่างทดลองเครื่องมือกลุ่มเดียวแล้วหาค่าความเชื่อมั่นได้

ข้อตกลงเบื้องต้นของวิธีนี้ คือ เครื่องมือชุดนั้นต้องวัดลักษณะเดียวกัน และมีระบบการให้คะแนนที่เป็น Dichotomous คือ ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน การหาความเชื่อมั่นวิธีนี้เป็นการหาความคงตัวภายใน (Internal consistency) มีสูตรในการคำนวณหาค่าความเชื่อมั่นอยู่ 2 สูตร คือ สูตร K.R. 20 และ สูตร K.R. 21

จ. แบบของครอนบาค (Cronbach) ในกรณีที่เครื่องมือที่สร้างให้คะแนนเป็นแบบจัดอันดับ หรือมาตราส่วนประมาณค่า เช่น ข้อสอบอัตนัย แบบสอบถาม มาตราวัดทัศนคติต่างๆ ครอนบาคเสนอแนะ ให้ใช้การหาความเชื่อมั่นโดยหาค่าสัมประสิทธิ์แอลฟา (α - Coefficient)

ฉ. แบบวิเคราะห์ความแปรปรวน (Analysis of variance) ซึ่งเป็นวิธีที่ฮอยท์ (Hoyt) เป็นผู้คิดขึ้น เป็นวิธีที่ใช้กับเครื่องมือที่ระบุการให้คะแนนไม่เป็น Dichotomous เช่น การสัมภาษณ์ ซึ่งมีผู้สัมภาษณ์และผู้ถูกสัมภาษณ์หลายคน

ช. การใช้วิธีอื่นๆ ในกรณีที่เครื่องมือวิจัยเป็นการสังเกต การหาความเชื่อมั่นที่นิยมกระทำกันใช้วิธีของวิลเลียม เอ สกอตต์ (William A. Scott)

บุญเชิด ภิญโญนนตพงษ์ (2537 : 12 - 46) ได้สรุปแนวคิดในการหาความเชื่อมั่นของ แบบทดสอบไว้ 2 แนวคิด คือ

1. การวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ ของการวัดหลายๆ ครั้ง จากการทดสอบซ้ำ ด้วยแบบทดสอบฟอร์มเดียวกันหรือแบบทดสอบฟอร์มที่คู่ขนานกัน ซึ่งมีค่าเท่ากับความแปรปรวนของคะแนนจริงหารด้วยความแปรปรวนของคะแนนที่สอบได้ โดยเรียกสัมประสิทธิ์สหสัมพันธ์ ที่คำนวณจากคะแนนที่ได้จากการสอบด้วยข้อสอบฉบับเดียวกันสองครั้งในเวลาแตกต่างกันว่า สัมประสิทธิ์ของความคงตัว (Coefficient of Stability) และเรียกสัมประสิทธิ์สหสัมพันธ์ ที่คำนวณจากคะแนนที่ได้จากการสอบด้วย ข้อสอบคู่ขนาน (Parallel Test) หรือแบบสลับฟอร์ม (Alternate Form) ที่สอบติดต่อกันที่ หรือสอบทั้งช่วงเวลาก็ได้ว่าสัมประสิทธิ์ของความสมมูล (Coefficient of Equivalent)

2. ใช้แบบทดสอบเพียงฉบับเดียวและสอบเพียงครั้งเดียว แล้วคำนวณหา สัมประสิทธิ์ความสอดคล้องภายในของแบบทดสอบ (Coefficient of Internal Consistency) จากความแปรปรวนของคะแนนสอบภายในฉบับจากส่วนย่อยแต่ละตอนและส่วนรวมทั้งฉบับ

นักทฤษฎีทางการวัดผลได้เสนอเทคนิค ในการประมาณค่าความเชื่อมั่นที่เชื่อถือได้ไว้หลายเทคนิค โดยมีข้อสันนิษฐานเบื้องต้น (presumptions) ว่า แบบทดสอบฉบับรวมสามารถแบ่งเป็นส่วนๆ เช่น สองส่วน สามส่วน สี่ส่วน หรือหลายๆ ส่วน และเมื่อใช้ระดับของความคู่ขนานของการวัด (degree of measurement parallelism) ในแต่ละส่วนเป็นเกณฑ์

แล้ว การประมาณค่าความเชื่อมั่นที่ได้จากการสอบเพียงครั้งเดียวด้วยข้อสอบเพียงฉบับเดียว จะจัดกลุ่มตามข้อตกลงของระดับความคู่ขนานได้ 3 กลุ่ม ดังนี้

แบบจำลองความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts)

เป็นการประมาณค่าความเชื่อมั่นจากแบบทดสอบแต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิม ที่มีข้อตกลงอย่างเคร่งครัด 6 ข้อ (บุญเชิด ภิญโญนนิตพงษ์. 2537 : 9 – 10) คือ

1. มีความเป็นเอกพันธ์ในเนื้อหา หรือวัดคุณลักษณะเดียวกัน
2. มีคะแนนจริงเท่ากัน ($T_{i1} = T_{i2} = \dots$) และมีความแปรปรวนคลาดเคลื่อนเท่ากัน ($S_{E1} = S_{E2} = S_{E3} = \dots$)
3. มีคะแนนสอบ (X) เฉลี่ยเท่ากัน ($\mu_1 = \mu_2 = \mu_3 = \dots$)
4. มีความแปรปรวนของคะแนนสอบ (X) เท่ากัน ($S^2_1 = S^2_2 = S^2_3 = \dots$)
5. มีความแปรปรวนร่วมของคะแนนสอบ (X) กับคะแนนสอบส่วนอื่นๆ

$$(S_{12} = S_{13} = S_{23} = \dots)$$

6. มีความแปรปรวนร่วมของคะแนนสอบ (X) กับเกณฑ์ภายนอกเท่ากัน

$$(S_{1y} = S_{2y} = S_{3y} = \dots)$$

จากข้อตกลงดังกล่าว ในทางปฏิบัติแทบไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิมได้ จึงได้มีการพัฒนาเทคนิคที่เหมาะสมขึ้นมาใหม่ โดยผ่อนปรนเงื่อนไขเดิมมาเป็น แต่ละส่วนจำเป็นต้องมีคะแนนจริงสมมูล (Essentially Tau – Equivalent Parts)

นักทฤษฎีการทดสอบที่มีชื่อเสียงได้แก่ สเปียร์แมน (Spearman. 1910) และบราวน์ (Brown. 1910)

แบบจำลองคะแนนจริงจำเป็นต้องสมมูล (Essentially Tau – Equivalent Parts)

วิธีนี้เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบ โดยผ่อนปรนเงื่อนไขข้อ 2, 3 และ 4 ของแบบจำลองความคู่ขนานแบบมาตรฐานเดิม ให้มีความเป็นไปได้มากขึ้น (บุญเชิด ภิญโญนนิตพงษ์. 2537 : 10 – 11) ดังนี้

- คะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากัน แต่ยอมให้ต่างกันได้เท่ากับ ความยากที่ต่างกันในแต่ละส่วน นั่นคือ $T_{ig} = T_{ih} + C_{gh}$ เมื่อ $g = h = 1, \dots, k$ และ C_{gh} ไม่จำเป็นต้องเท่ากับศูนย์เสมอไป

- แต่ละส่วนมีคะแนนสอบ (X) เฉลี่ยต่างกัน
- ความแปรปรวนของคะแนนสอบ (X) ต่างกันเล็กน้อย

แม้ว่าแต่ละส่วนของแบบทดสอบที่แบ่งนั้นจะมีขนาดของความยาวหรือจำนวนข้อเท่ากัน แต่ในทางปฏิบัติจริงมีแบบทดสอบบางชนิดอาจต้องแบ่งส่วนให้เหมาะสมตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีจำนวนข้อไม่เท่ากัน ซึ่งส่งผลต่อเงื่อนไขข้อ 5 และ 6 ถึงแม้ว่าขนาดของความยาวเท่ากัน แต่เมื่อนำไปสอบกับกลุ่มตัวอย่างแล้วปรากฏว่า แต่ละส่วนมีการกระจายของคะแนนไม่เท่ากัน แสดงว่า ความยาวที่ทำหน้าที่ (Functional lengths) ของมันในแต่ละส่วนมีขนาดไม่เท่ากัน (Feldt. 1975 : 558) นักทฤษฎีการทดสอบจึงจำเป็นต้องนิยามความคู่ขนานขึ้นมาอีกระดับหนึ่ง ที่มีความผ่อนปรนมากที่สุด เรียกว่า ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model)

นักทฤษฎีที่เสนอเทคนิคสำหรับการประมาณค่าความเชื่อมั่นตามกลุ่มนี้ได้แก่ ฟลานาแกน (Flanagan. 1930) รูลอน (Rulon. 1930) กัตต์แมน (Guttman. 1945) เฟลด์และเบรนนอน (Feldt and Brennan. 1989) ครอนบาค (Cronbach. 1951)

ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model)

เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบที่ผ่อนปรนเงื่อนไขข้อ 1 ที่ว่า แต่ละส่วนของแบบทดสอบต้องมีเนื้อหาที่เป็นเอกพันธ์ หรือวัดคุณลักษณะ เดียวกัน (Kristof. 1974 ; 492)

ลักษณะสำคัญของความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ คือ การแบ่งแบบทดสอบออกเป็นส่วนๆ ที่มีขนาดความยาวไม่เท่ากัน หรือมีขนาดความยาวเท่ากัน แต่ก็มีการกระจายของคะแนนในแต่ละส่วนแตกต่างกันมาก

ซึ่งวิธีการคำนวณความเชื่อมั่น โดยใช้ระดับความคู่ขนานเป็นเกณฑ์กับแบบทดสอบที่สามารถแบ่งได้เป็นส่วนนั้น สามารถสรุปได้ดังนี้ (บุญเชิด ภิญโญอนันตพงษ์. 2545 : 216 – 230)

1) วิธีหาค่าความเชื่อมั่นแบบแบ่งหลายส่วนย่อยที่มีความคู่ขนานแบบมาตรฐานเดิม กรณีที่หลายส่วนมีความคู่ขนานแบบมาตรฐานเดิม ใช้สูตรของ สเปียร์แมน – บราวน์ (Spearman- Brown) สูตรทั่วไป ดังนี้

$$r_{tt} = \frac{nr_{pp}}{1 + (n-1)r_{pp}}$$

เมื่อ	r_{tt}	แทน	ความเชื่อมั่นของเครื่องวัดทั้งฉบับ
	r_{pp}	แทน	ความเชื่อมั่นส่วนย่อย
	n	แทน	จำนวนส่วนย่อยในฉบับ

2) วิธีหาค่าความเชื่อมั่นแบบแบ่งหลายส่วนย่อยที่มีความคู่ขนานแบบคะแนนจริงสมมูลถ้าส่วนย่อยของเครื่องมือวัดแบ่งเป็นหลายส่วนและมีคู่ขนานแบบคะแนนจริงสมมูล แยกพิจารณาเป็นอีกสองกรณีคือ

2.1) เครื่องมือวัดให้คะแนนแบบสองค่า เช่น แบบทดสอบเลือกตอบ แบบทดสอบตอบถูกผิดหรือแบบทดสอบเติมคำ เมื่อตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน หรือแบบสอบถามบางเรื่องที่ให้เลือกตอบว่าใช่ หรือไม่ใช่ ถ้าใช่ให้ 1 คะแนน ถ้าไม่ใช่ให้ 0 คะแนน การหาค่าความเชื่อมั่นคำนวณได้จากสูตร คูเดอร์ – ริชาร์ดสัน (Kuder – Richardson) สูตร 20 และ 21

สูตร KR – 20 ใช้สำหรับคำนวณค่าความเชื่อมั่นของเครื่องมือวัดที่ให้คะแนนแบบสองค่าเป็น 0 หรือ 1 มีสูตรดังนี้

$$r_{tt} = \frac{K}{K-1} \left[1 - \frac{\sum pq}{S_x^2} \right]$$

เมื่อ	r_{tt}	แทน	ค่าความเชื่อมั่นของเครื่องมือวัด
	K	แทน	จำนวนข้อของเครื่องมือวัด
	p	แทน	สัดส่วนของผู้ตอบถูกหรือความยากของแต่ละข้อ
	q	แทน	สัดส่วนของผู้ตอบผิด ซึ่งเท่ากับ $1 - P$
	S_x^2	แทน	ความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

สูตร KR – 21 เป็นสูตรที่ปรับมาจากสูตร KR – 20 โดยสมมติว่าข้อสอบแต่ละข้อมีความยากเท่ากัน มีสูตรดังนี้

$$r_{tt} = \frac{K}{K-1} \left[1 - \frac{\bar{X}(K - \bar{X})}{KS_x^2} \right]$$

เมื่อ	r_{tt}	แทน	ค่าความเชื่อมั่น
	K	แทน	จำนวนข้อของเครื่องมือวัด
	$\bar{X}$	แทน	คะแนนเฉลี่ยของผลการทดสอบ
	S_x^2	แทน	ความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

ข้อแตกต่างระหว่างสูตร KR – 20 และสูตร KR – 21 คือการหาค่าความเชื่อมั่นของสูตร KR – 21 มีข้อสมมติเพิ่มว่า ข้อคำถามแต่ละข้อต้องมีความยากเท่ากัน ดังนั้นไม่ต้องหาค่าความยากรายข้อ ทำให้เกิดการคิดคำนวณง่ายขึ้นและหาได้รวดเร็ว เพราะเพียงแต่ค่าเฉลี่ยของค่าเฉลี่ยของกลุ่ม และค่าความแปรปรวน ของคะแนนรวมทั้งฉบับ ซึ่งเป็นค่าสถิติของเครื่องมือวัด ที่ต้องคำนวณไว้ก่อนแล้ว ทำให้ค่าความเชื่อมั่นที่คำนวณด้วยสูตร KR – 21 ได้ค่าประมาณที่ต่ำกว่าความเป็นจริงเสมอ ดังนั้นผลการคำนวณจากสูตร KR – 21 จะได้ค่าน้อยกว่าที่คำนวณจากสูตร KR – 20 เสมอ

2.2) ถ้าเครื่องมือวัดนั้นให้คะแนนหลายค่า เช่น มาตรฐานประเมินค่า (Rating Scale) บุคลิกภาพฉบับหนึ่ง ให้คะแนนเป็นหลายค่า เมื่อข้อความนั้นตรงตามลักษณะของผู้ที่ตอบจากน้อยสุดไปหามากสุด จะได้คะแนนเป็น 1, 2, 3, 4 และ 5 หรือแบบทดสอบอัตนัยฉบับหนึ่งมีคำถาม 5 ข้อ แต่ละข้อคะแนนเต็ม 10 คะแนน ดังนั้นแต่ละข้อแต่ละข้อสามารถให้คะแนนหลายค่าตั้งแต่ 0 ถึง 10 คะแนน ในกรณีนี้ การหาค่าความเชื่อมั่นคำนวณได้จากสูตร สัมประสิทธิ์แอลฟา (Coefficient - α) หรือสูตรของครอนบัค (Cronbach)

สูตรสัมประสิทธิ์แอลฟา (Coefficient - α) ของครอนบัคมีสูตร ดังนี้

$$r_{tt} \text{ หรือ } \alpha = \frac{K}{K-1} \left[1 - \frac{\sum S_i^2}{S_x^2} \right]$$

เมื่อ α	แทน	ค่าความเชื่อมั่นของเครื่องมือวัด
K	แทน	จำนวนส่วนย่อยหรือจำนวนข้อคำถามของเครื่องมือวัด
S_i^2	แทน	ค่าความแปรปรวนของคะแนนแต่ละส่วนย่อยหรือแต่ละข้อ
S_x^2	แทน	ค่าความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

3) วิธีหาค่าความเชื่อมั่นกรณีหลายส่วนมีความคู่ขนานแบบคะแนนจริงสัมพัทธ์

3.1) เมื่อใช้จำนวนข้อเป็นตัวกำหนดความยาว (Nominal Lengths) ของเครื่องมือวัด โดยแบ่งเครื่องมือวัดเป็นส่วนๆ เท่ากับจำนวนข้อ หรือแต่ละส่วนมีความยาวเท่ากับ 1 ข้อ และแต่ละข้อให้คะแนนเป็นสองค่า คือ ตอบถูก ได้ 1 ตอบผิด ได้ 0 สามารถคำนวณจากสัมประสิทธิ์ r_B (Coefficient r_B) ของ บัญเชิด ภิญโญอนันตพงษ์ (2537 : 51) ซึ่งสามารถพัฒนามาจากสัมประสิทธิ์ r_{12} ของ Liou มีสูตรดังนี้

สัมประสิทธิ์ r_B

$$r_B = \frac{1}{1 - \sum \lambda_i^2} \left[1 - \frac{\sum P_i (1 - P_i)}{S_x^2} \right]$$

เมื่อ r_B แทน ความเชื่อมั่นของแบบทดสอบ
 P_i แทน ค่าความยากของข้อสอบแต่ละข้อ
 S_x^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ

ค่า λ_i หาได้จากสูตรต่อไปนี้

$$\lambda_i = [\sum B_{1i} - P_1 (\bar{X} - 1)] / S_x^2$$

$$\lambda_i = [\sum B_{2i} - P_2 (\bar{X} - 1)] / S_x^2$$

$\vdots$

$$\lambda_i = [\sum B_{ki} - P_k (\bar{X} - 1)] / S_x^2$$

เมื่อ $\bar{X}$ แทน ผลรวมของค่าความยากแต่ละข้อ
 B_{ji} แทน สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใดๆ ถูกทั้งสองข้อ

สูตร r_B คล้ายกับสูตร KR - 20 แต่ใช้ข้อตกลงแบบคะแนนจริงสัมพัทธ์
 ดังนั้นสูตร r_B จึงเป็นสูตรปรับขยายสูตร KR - 20

3.2) เมื่อใช้จำนวนข้อเป็นตัวกำหนดความยาว (Nominal Lengths) ของ
 เครื่องมือวัดคำนวณได้จากสูตร ราชู (Raju) สูตรนี้ใช้หาค่าความเชื่อมั่นเมื่อแบ่งเครื่องมือวัด
 ออกเป็นหลายๆ ส่วน แต่ละส่วนมีจำนวนข้อหรือความยาวไม่เท่ากัน ราชู ตั้งชื่อสูตรนี้ว่า
 สัมประสิทธิ์เบต้าเค (Coefficient - β_K) ถ้าแต่ละส่วนมีจำนวนข้อเท่ากันแล้ว ค่าสัมประสิทธิ์เบต้าเคจะ
 เท่ากับสัมประสิทธิ์แอลฟาของครอนบัค ดังนั้นสัมประสิทธิ์ เบต้าเคจึงเป็นสูตรทั่วไปของ
 สัมประสิทธิ์แอลฟา มีสูตรดังนี้

$$r_{tt} \text{ หรือ } \beta_K = \left[\frac{1}{1 - \sum p_i^2} \right] \left[1 - \frac{\sum S_i^2}{S_x^2} \right]$$

$$p_i = \frac{n_i}{\sum n_i}$$

เมื่อ β_K	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
p_i	แทน	สัดส่วนของจำนวนข้อแต่ละส่วนย่อย
S_i^2	แทน	ความแปรปรวนของคะแนนในแต่ละส่วน
S_x^2	แทน	ความแปรปรวนของคะแนนทั้งฉบับ
n_i	แทน	จำนวนข้อในแต่ละส่วนย่อยของแบบทดสอบ
$\sum n_i$	แทน	จำนวนข้อสอบทั้งหมด

3.3) เมื่อใช้ผลการสอบจริงเป็นตัวชี้บอกความยาว (Functional of Effective Lengths) ของ เครื่องมือวัด หรือ ไม่ทราบความยาว (Unknown Lengths) จำนวนได้จากสูตร เฟลด์ - ราจู (Feldt - Raju) สูตรทั่วไปที่แบ่งส่วนย่อยเป็น K ส่วน ดังนี้

สูตร เฟลด์ - ราจู (Feldt - Raju) สูตรทั่วไป

$$r_{tt} = \left[\frac{1}{1 - \sum \lambda_i^2} \right] \left[1 - \frac{\sum S_i^2}{S_x^2} \right]$$

$$\lambda_1 = \frac{\sum S_{1i}}{S_x^2}$$

$$\lambda_K = \frac{\sum S_{Ki}}{S_x^2}$$

เมื่อ r_{tt}	แทน	ค่าความแปรปรวนของเครื่องมือวัดหรือแบบทดสอบ
S_i^2	แทน	ความแปรปรวนของคะแนนในแต่ละตอน
S_x^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ

ปัจจัยที่ส่งผลกระทบต่อค่าความเชื่อมั่น (บุญเชิด ภิญโญนนตพงษ์. 2545 : 237 – 241 และ สุมาลี จันทรชลอ. 2542 : 110 – 113)

1. ความเป็นเอกพันธ์ของกลุ่มตัวอย่าง (Group homogeneity) สัมประสิทธิ์ความเชื่อมั่นเป็นค่าที่ได้จากการวัดในแต่ละครั้งว่า สามารถวัดลำดับที่ของคนในกลุ่มหนึ่งๆ ได้คงเส้นคงวาเพียงใด ถ้าลักษณะของคำตอบคล้ายคลึงกันมาก หรือมีความเป็นเอกพันธ์เนื่องจากผู้สอบความสามารถใกล้เคียงกัน ความแปรปรวนต่ำ ค่าความเชื่อมั่นจะลดลง ดังนั้นโอกาสที่จะทำให้ลำดับที่ของการวัดคลาดเคลื่อนไปจากเดิมจะมีมากขึ้น แต่ถ้าลักษณะคำตอบของกลุ่มตัวอย่างมีความแตกต่างกันมาก (Heterogeneity) กล่าวคือ มีผู้สอบเก่งมาก และอ่อนมากๆ สอบด้วยกันหรือเป็นกลุ่มวิวิธพันธ์ ค่าความแปรปรวนจะสูงขึ้น ค่าความเชื่อมั่นจะสูงขึ้น ดังนั้น ผลที่ได้จากการวัดจะมีลำดับที่คงเดิมมาก

ในกรณีทราบค่าความแปรปรวนของคะแนนกลุ่มที่ต้องการทดสอบ ผู้ใช้เครื่องมือวัดหาค่าความเชื่อมั่นได้โดยใช้สูตรดังนี้

$$r_{nn} = 1 - \frac{S_o^2}{S_n^2} [1 - r_{oo}]$$

เมื่อ	r_{nn}	แทน	ความเชื่อมั่นของเครื่องมือวัดที่สอบกับนักเรียนกลุ่มใหม่
	r_{oo}	แทน	ความเชื่อมั่นของเครื่องมือวัดที่สอบกับนักเรียนกลุ่มเดิม
	S_n^2	แทน	ความแปรปรวนของนักเรียนกลุ่มใหม่
	S_o^2	แทน	ความแปรปรวนของนักเรียนกลุ่มเดิม

2. ความยาวของเครื่องมือวัดความยาวหรือจำนวนของข้อของเครื่องมือวัด มีผลกระทบต่อค่าความเชื่อมั่น โดยที่เครื่องมือวัดใดมีจำนวนข้อน้อย จะมีค่าความเชื่อมั่นต่ำกว่า ถ้ามีจำนวนข้อมาก สามารถวัดพฤติกรรมได้ครอบคลุมและมีมากพอ คะแนนที่ได้มีการเดาน้อยลง ถ้าองค์ประกอบอื่นคงที่ สูตรที่แสดงให้เห็นความสัมพันธ์ระหว่างความยาวของแบบทดสอบกับความเชื่อมั่นคือ สูตรทั่วไปของสเปียร์แมน – บราวน์ (Spearman – Brown) ที่ปรับมาจากสูตรดังนี้

$$r_{nn} = \frac{nr_{oo}}{1 + (n-1)r_{oo}}$$

เมื่อ	r_{nn}	แทน	ความเชื่อมั่นของเครื่องมือวัดฉบับใหม่
	r_{oo}	แทน	ความเชื่อมั่นของเครื่องมือวัดฉบับเดิม
	n	แทน	จำนวนเท่าที่เพิ่มหรือลดลง

3. ระดับความสามารถของนักเรียนในกลุ่ม สิ่งที่มีอิทธิพลต่อค่าความเชื่อมั่นที่เกิดจากตัวนักเรียนเองอีกแบบหนึ่ง คือ ระดับความสามารถของนักเรียน ถ้านักเรียนในกลุ่มมีความสามารถเฉลี่ยสูงจะสามารถทำข้อสอบได้ถูกต้องแน่นอนทุกครั้ง แต่ถ้านักเรียนในกลุ่มมีความสามารถเฉลี่ยต่ำ มักจะตอบโดยการเดาเป็นส่วนใหญ่ จึงทำให้คะแนนจากการสอบไม่แน่นอน ซึ่งทำให้ค่าความเชื่อมั่นต่ำลง

4. ระดับความยากง่าย ข้อความยากของข้อสอบมีอิทธิพลต่อค่าความเชื่อมั่นในแง่ที่ทำให้การกระจายของคะแนนมีมากน้อยต่างกัน ข้อสอบที่ยากเกินไปจะมีค่าความแปรปรวนต่ำ เพราะการกระจายของคะแนนข้อนั้นๆ มีน้อย แต่ถ้าข้อสอบมีความยากง่ายพอเหมาะ ช่วงการกระจายของคะแนนกว้าง ค่าความเชื่อมั่นจะสูง

5. ความคล้ายคลึงของเนื้อหาที่ออกข้อสอบ ข้อสอบที่วัดลักษณะเดียวกันทั้งฉบับหรือมีเนื้อหา ของข้อสอบคล้ายคลึงกัน จะมีค่าความเชื่อมั่นสูงกว่าข้อสอบที่มีเนื้อหาที่แตกต่างกันมากๆ เช่น ข้อสอบที่ออกเนื้อหาเพียงบทเดียว ซึ่งมีเนื้อหาเดียว จะมีแนวโน้มให้ค่าความเชื่อมั่นสูงกว่าข้อสอบที่ออกเนื้อหาแตกต่างกันหลายๆ บทหรือทั้งเล่ม

6. แบบทดสอบแบบเร่งรีบ (Speed test) แบบทดสอบประเภทนี้ประกอบด้วยข้อสอบที่ง่าย ๆ แต่มีมากข้อ ผู้ตอบต้องอาศัยความเร็วในการตอบคำถาม ซึ่งส่วนมากแล้วนักเรียนสามารถตอบถูก ทุกข้อที่ทัน นั่นคือ ตอบถึงข้อใดก็มักจะได้คะแนนเท่านั้นเสมอ ดังนั้นการสอบแต่ละครั้งจึงมีคะแนนคงเส้นคงวาเหมือนเดิมเสมอ ทำให้เครื่องมือวัดประเภทนี้มีค่าความเชื่อมั่นค่อนข้างสูง

7. วิธีการคำนวณ (Computed method) การใช้สูตรในการคำนวณแตกต่างกัน จะให้ค่าความเชื่อมั่นต่างกัน เช่น KR_{20} มักให้ค่าความเชื่อมั่นสูงกว่า KR_{21} การคำนวณ โดยใช้ Split – Half method อาจให้ค่าความเชื่อมั่นสูง แต่ไม่เหมาะสมกับข้อสอบที่จำกัดเวลา เพราะผู้สอบต้องให้ความเร็ว (Speed test) การเลือกใช้วิธีการคำนวณ ต้องคำนึงถึงองค์ประกอบอื่นๆ เช่น วิธีการให้คะแนน หรือวิธีการสอบวัด แบบทดสอบที่มี วิธีวัดให้คะแนนเป็น 0, 1 การคำนวณค่าความเชื่อมั่นจะใช้สูตรคูเดอร์ ริชาร์ดสัน (Kuder Richardson) KR_{20} , KR_{21} ถ้าแบบทดสอบมีลักษณะการให้คะแนนไม่เป็น 0, 1 การคำนวณค่าความเชื่อมั่นอาจใช้สูตร อัลฟา (α) ของครอนบัค ถ้าเป็นการสอบวัดด้วยแบบทดสอบคู่หรือแบบทดสอบฉบับเดียวกัน 2 ครั้ง ในระยะเวลาที่ห่างกันพอประมาณ อาจคำนวณค่าความเชื่อมั่นโดยใช้สูตรของเพียร์สัน (Pearson Product Moment) เป็นต้น วิธีการคำนวณดังกล่าวให้ค่าความเชื่อมั่นต่างกัน เช่น KR_{20} อาจให้ค่าความเชื่อมั่นสูงกว่า KR_{21} ดังนั้น การตีความค่าความเชื่อมั่น ควรคำนึงถึงสูตรในการคำนวณด้วย

8. เวลาในการทำแบบทดสอบ การทดสอบผู้เรียนควรให้เวลาในการทำข้อสอบที่เหมาะสม ถ้าให้เวลาน้อยไป ผู้สอบทำไม่ทันอาจเดาคำตอบ ทำให้ผลที่ได้แตกต่างไปจากความสามารถที่แท้จริง ซึ่งจะแตกต่างไปจากการทดสอบเมื่อให้เวลาพอเหมาะถ้าให้เวลามากเกินไป ผู้เรียนอ่อนมีเวลาคิดใคร่ครวญ ทำให้ได้คะแนนดีแตกต่างกันไป จากความสามารถที่แท้จริง เมื่อทำการสอบอีกครั้งโดยให้เวลาพอเหมาะ ผลของคะแนนจะไม่คงที่ และเมื่อพิจารณาในแง่ของความแปรปรวน เมื่อให้เวลาน้อย ผู้สอบทำไม่ทันและได้คะแนนน้อยเหมือนกัน หรือเมื่อให้เวลามาก ผู้สอบส่วนมากทำได้ใกล้เคียงกัน คะแนนจะมีความแปรปรวนน้อย ซึ่งมีผลทำให้ความเชื่อมั่นลดลง

9. ตัวอย่างประชากรที่ใช้ทดลอง จะมีผลกระทบต่ค่าความเชื่อมั่นของเครื่องมือวัด ถ้ากลุ่มตัวอย่าง ที่นำมาทดลองมีจำนวนน้อยเกินไปหรือในบางครั้งไม่เป็นตัวแทนของประชากร เช่น สร้างแบบทดสอบวัดความสามารถในการคิดคำนวณระดับวิทยาลัยสำหรับนิสิตสาขาวิทยาศาสตร์ แต่เวลาทดลองเครื่องมือกลับไปสุ่มนักเรียนสาขากฎหมายหรือสาขาสังคมมาทดลองหาค่าความเชื่อมั่น ทำให้ค่าความเชื่อมั่นของเครื่องมือวัดที่คำนวณได้ขาดความเชื่อถือ นอกจากนั้นความไม่คุ้นเคยกับแบบทดสอบ อารมณ์ การเจ็บป่วย ความวิตกกังวล สิ่งเหล่านี้มีผลกระทบต่ค่าความเชื่อมั่นทั้งสิ้น

10. ความเป็นปรนัย (Objectivity) ในที่นี้หมายถึงความกระจัดกระจาย แม่นยำ และปลอดจากความคลาดเคลื่อน เครื่องมือวัดที่มีความเป็นปรนัยสูงจึงเป็นเครื่องมือที่มีลักษณะสำคัญ คือ 1) ความแจ่มชัดในตัวเครื่องมือวัด คำอธิบายหรือคำชี้แจงในการตอบคำถามตรงกันและตรงกับ ความมุ่งหมายของเครื่องมือวัด คำอธิบายหรือคำชี้แจงในการตอบชัดเจน รู้ชัดว่าต้องการถามอะไรไม่กำกวม ไม่วกวน ไม่พลิกแพลง เล่นภาษา 2) ความชัดเจนในการตรวจให้คะแนนมี เกณฑ์ในการให้คะแนนไว้กระจัดกระจาย มีกฎเกณฑ์ที่แน่นอนว่าการตอบถูกอย่างไรจะได้คะแนนเท่าไร ไม่ว่าใครมาตรวจจะได้คะแนนตรงกันเพราะมีเกณฑ์ในการให้คะแนนอย่างเดียวกัน เช่น ตอบถูกได้ 1 คะแนน ผิดได้ 0 3) มีความกระจัดกระจายในการตีความหมายคะแนน เช่น คะแนน แทนความรู้ที่ทำได้สูงกว่าเป็นผู้มีความสามารถจริงสูงกว่า ได้คะแนนน้อย กล่าวคือ มีอำนาจจำแนกสูง คะแนนที่ได้ขึ้นอยู่กับความยากของเครื่องมือวัดด้วย ดังนั้น การแปลความหมายของคะแนนแบบอิงกลุ่ม จึงต้องเปรียบเทียบอันดับภายในกลุ่มมากกว่าที่จะพึ่งเล็งคะแนนดิบ คะแนนอิงเกณฑ์ก็ต้องเป็นตัวแทนที่จะเป็นดัชนีบ่งชี้คุณลักษณะของบุคคลในเรื่องที่วัด กล่าวคือ เครื่องมือวัดเป็นตัวแทนกลุ่มงาน กลุ่มของพฤติกรรมย่อย (Domain of behavior) ดังนั้น คะแนนที่เป็นปรนัยในแนวอิงเกณฑ์จึงเป็นคะแนนที่บ่งระดับของคุณลักษณะหรือความสามารถในเรื่องที่เครื่องมือวัดนั้นวัด

เครื่องมือวัดที่มีความเป็นปรนัยสูงจึงมีความคลาดเคลื่อนต่างๆ ปะปนอยู่น้อย ส่งผลให้คะแนนที่ได้ใกล้เคียงกับคะแนนจริงซึ่งเป็นเงื่อนไขสำคัญอย่างหนึ่งที่จะทำให้มีความเชื่อมั่นสูง ข้อสอบอัตนัยมีความเป็นปรนัยต่ำ จึงมีค่าความเชื่อมั่นต่ำ

ความคลาดเคลื่อนมาตรฐานของการวัด (บุญเชิด ภิญโญนนัดพงษ์. 2545 : 248)

ความคลาดเคลื่อนมาตรฐานของการวัด มีความสัมพันธ์กับค่าความเชื่อมั่น ถ้าเครื่องมือวัดมีความเชื่อมั่นสูง ($r_e = 1.00$) ความคลาดเคลื่อนในการวัดจะไม่มี ถ้าความเชื่อมั่นของเครื่องมือวัดต่ำแล้วค่าความคลาดเคลื่อนมาตรฐานในการวัดจะสูง จากนิยามของความเชื่อมั่นจะเห็นว่าความเชื่อมั่นจะสูงขึ้นเมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนลดลงและมีความแปรปรวนของคะแนนสอบคงที่หรือเมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนคงที่และความแปรปรวนของคะแนนสอบเพิ่มขึ้น ความเชื่อมั่นจะสูงขึ้นเช่นกัน เนื่องจาก และ $r_e = 1 - \frac{S_e^2}{S_x^2}$ ค่ารากที่สองของความแปรปรวนคะแนน $S_e^2 = S_x^2(1 - r_e)$ ความคลาดเคลื่อนคือความคลาดเคลื่อนมาตรฐานหรือ $S_e = S_x \sqrt{1 - r_e}$

ซึ่งจากที่กล่าวมาข้างต้นสามารถสรุปได้ว่า ความเชื่อมั่น หมายถึง ความคงที่แน่นอนของคะแนน ที่ได้จากการสอบวัดจากเด็กกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวหรือขนานกัน ในเวลาต่างกัน หรืออาจกล่าวได้ว่า ความเชื่อมั่นเท่ากับสัดส่วนของความแปรปรวนที่ได้จากคะแนนจริงต่อความแปรปรวนของคะแนนที่สอบได้ โดยที่ความเชื่อมั่นของการวัดอาจแบ่งเป็นชนิดต่างๆ หลายแบบ ถ้าแบ่งตามลักษณะการสอบวัด สามารถจำแนกสัมประสิทธิ์ของการคำนวณได้ 3 ชนิด คือ 1) สัมประสิทธิ์ความคงตัว 2) สัมประสิทธิ์ความสมมูล และ 3) สัมประสิทธิ์ของความสอดคล้องถ้าแบ่งความเชื่อมั่นของเครื่องมือวัดแบบอิงกลุ่มและแบบอิงเกณฑ์ 1) ความเชื่อมั่นของเครื่องมือวัดแบบอิงกลุ่ม อาจเป็นแบบวัดซ้ำ แบบคู่ขนาน และแบบแบ่งส่วนภายในฉบับ ซึ่งอาจจะแบ่งเป็น 2 ส่วน หรือหลายส่วน 2) ความเชื่อมั่นของการวัดแบบอิงเกณฑ์ แบ่งเป็นวิธีประมาณค่าความเชื่อมั่นในการตัดสินจำแนกความรอบรู้และวิธีประมาณค่าความเชื่อมั่นของคะแนนแบบทดสอบอิงเกณฑ์ ถ้าแบ่งตามระดับของความคู่ขนาน อาจจำแนกเป็น 3 ชนิด คือ 1) ความเชื่อมั่นแบบคะแนนจริงมาตรฐานเดิม ซึ่งนักทฤษฎีการทดสอบที่มีชื่อเสียงได้แก่ สเปียร์แมน (Spearman) และ บราวน์ (Brown) เป็นต้น 2) ความเชื่อมั่นแบบคะแนนจริงสมมูล ซึ่งนักทฤษฎีที่เสนอเทคนิคสำหรับการประเมินค่าความเชื่อมั่นตามกลุ่มนี้ ได้แก่ ฟลานาแกน (Flanagan) รูลอน (Rulon) กัตต์แมน (Guttman) เฟลด์ และเบรนนอน (Feldt and Brennan) และครอนบัค (Cronbach) เป็นต้น 3) ความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ ซึ่งลักษณะสำคัญของความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์คือการแบ่งแบบทดสอบออกเป็นส่วนๆ ที่มีขนาดความยาวไม่เท่ากัน หรือมีขนาดความยาวเท่ากัน แต่มีการกระจายของคะแนนในแต่ละส่วนแตกต่างกันมาก ซึ่งนักทฤษฎีที่เสนอทฤษฎีสำหรับการประมาณค่าความเชื่อมั่นของกลุ่มนี้ ได้แก่ ราจู (Raju) เฟลด์ - ราจู (Feldt - Raju)

ฮอร์ส (Horst) เฟลด์ (Feldt) เลี้ยว (Liou) และคริสทอฟ (Kristof) เป็นต้น ซึ่งในที่นี้ ผู้วิจัยจึงมีความสนใจที่จะศึกษาค่าความเชื่อมั่นของมาตราประเมินค่าจากสูตรประสิทธิ์แอลฟา (α) ของครอนบัค เนื่องจากเป็นที่ที่นิยมและรู้จักกันอย่างแพร่หลาย แต่เนื่องจากการหาค่าความเชื่อมั่นแบบสัมประสิทธิ์แอลฟา (α) ของครอนบัคนั้น ได้วิเคราะห์ส่วนย่อยของแบบทดสอบเป็นส่วนย่อยๆ รายข้อเป็นไปตามข้อตกลงคะแนนจริงสมมูล แต่ในทางปฏิบัติในบางวิชาอาจไม่สามารถแบ่งแบบทดสอบเป็นรายข้อได้ เช่น แบบทดสอบวิชาภาษาอังกฤษ ซึ่งส่วนใหญ่แบบทดสอบในวิชานี้จะมีหลายตอน (Parts) ดังนั้น เราควรจะวิเคราะห์ส่วนย่อยออกเป็นรายตอนมากกว่า (Feldt & Brennan, 1989 : 112) ซึ่งวิธีการนี้ควรใช้แบบจำลองของคะแนนจริงสัมพันธ์มาทำการวิเคราะห์ซึ่งมีผู้เสนอสูตรอยู่หลายสูตรแต่ละสูตรจะแตกต่างกันในจำนวนส่วนย่อย ซึ่งผู้วิจัยตัดสินใจที่จะใช้สูตร r_{F-R} ของ เฟลด์ ราชู ด้วย เนื่องจากสูตรนี้เป็นไปตามข้อตกลงแบบจำลองคะแนนจริงสัมพันธ์ เมื่อแบ่งแบบทดสอบออกเป็น K ส่วน และเนื่องจากสูตรนี้ได้พัฒนามาจากสูตรของครอนบัค นั่นเอง

4. เอกสารและงานวิจัยที่เกี่ยวข้อง

4.1 งานวิจัยในประเทศไทย

นางบุญ เจนสรรพกิจกุล (2534 : บทคัดย่อ) ได้ศึกษาการเปรียบเทียบผลการตอบมาตราประเมินค่าที่มีรูปแบบคำตอบและลักษณะคำถามต่างกัน โดยเปรียบเทียบแบบแผนและความคงที่ในการตอบ มาตราประเมินค่าแบบลิเคอร์ท เมื่อกำหนดรูปแบบคำตอบและลักษณะคำถามแตกต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร จำนวน 400 คน ใช้แบบสอบถามความคิดเห็นเกี่ยวกับเรื่องเพศของอรรถ ปิตานนท์ เป็นเครื่องมือในการเก็บรวบรวมข้อมูล ผลการวิจัยพบว่า

1. ผู้ตอบมาตราประเมินค่าที่มีรูปแบบคำตอบแตกต่างกัน 5 รูปแบบ มีแบบแผนการตอบแตกต่างกัน แต่ไม่พบความแตกต่างในด้านความคงที่ในการตอบ
2. ผู้ตอบประเมินค่ามีลักษณะคำถามแตกต่างกันระหว่างคำถามที่วัดต่อความรู้สึกกับวัดต่อความรู้สึกรู้ มีแบบแผนการตอบแตกต่างกัน แต่ไม่พบความแตกต่างในด้านความคงที่ในการตอบ
3. ไม่พบปฏิสัมพันธ์ระหว่างลักษณะคำถามกับรูปแบบคำตอบต่อความคงที่ในการตอบ มาตราประเมินค่า

เพ็ญศรี ลาหุลำเลิศ (2534 : บทคัดย่อ) ได้เปรียบเทียบค่าสถิติของมาตรประเมินค่าที่มีรูปแบบคำตอบแตกต่างกัน 5 แบบคือ แบบกราฟิกอธิบายหัวข้อ แบบกราฟิกอธิบายทุกสเกล แบบบรรยาย-อย่างง่าย แบบบรรยายที่สอดคล้องพฤติกรรม และแบบเปอร์เซ็นต์ของพฤติกรรมในการประเมินตนเอง ของนักเรียนชั้นมัธยมศึกษาตอนปลาย ผลการวิจัยพบว่ามาตรประเมินค่าทั้ง 5 รูปแบบ มีค่าเฉลี่ย ความแปรปรวน ลักษณะการแจกแจงความถี่ ความเที่ยงตรง และความเชื่อมั่น แตกต่างอย่างไม่มีนัยสำคัญทางสถิติ นอกจากนี้ยังพบว่ามาตรประเมินค่าทั้ง 5 รูปแบบมีค่าความเชื่อมั่นสูงกว่า .80 แต่รูปแบบที่ 3, 4 และ 5 อันเป็นแบบบรรยายอย่างง่าย แบบบรรยายที่สอดคล้องพฤติกรรมและแบบเปอร์เซ็นต์ของ พฤติกรรม เป็นรูปแบบที่ดีกว่ารูปแบบกราฟิกที่อธิบายหัวข้อ และแบบกราฟิกที่อธิบายทุกสเกล ในด้านความเที่ยงตรงตามสภาพ

สุภาวดี ตั้งบุผา (2534 : บทคัดย่อ) ได้เปรียบเทียบความคลาดเคลื่อนของการประเมินค่าที่ใช้รูปแบบของมาตรประเมินค่าและผู้ประเมินที่แตกต่างกัน มาตรประเมินค่าที่ใช้มี 3 รูปแบบ คือ แบบ ตัวเลข แบบกราฟิก และแบบบรรยาย ผลการวิจัยพบว่า กรณีที่นิสิตเป็นผู้ประเมินคุณลักษณะอาจารย์ ที่ปรึกษากิจการนิสิต พบว่ามาตรประเมินค่าแบบตัวเลขและแบบบรรยายมีความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนนและมีความคลาดเคลื่อนแบบขัดแย้ง ในขณะที่แบบกราฟิกมีความคลาดเคลื่อนที่เกิดขึ้นทั้ง 3 ชนิดคือ ความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนน ความคลาดเคลื่อนจากปฏิสัมพันธ์ และความคลาดเคลื่อนแบบขัดแย้ง กรณีที่ผู้บริหารเป็นผู้ประเมิน จะพบว่าทั้ง 3 รูปแบบมีความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนน แต่ไม่มีความคลาดเคลื่อนอันเนื่องจากปฏิสัมพันธ์และความคลาดเคลื่อนแบบขัดแย้ง สำหรับการเปรียบเทียบความคลาดเคลื่อนของการประเมินค่าระหว่างมาตรประเมินค่าและผู้ประเมินพบว่า ผู้ประเมินที่เป็นนิสิตมีความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนนสูงกว่าผู้ประเมินที่เป็นผู้บริหาร

ศิรินาถ สิงหาแก้ว (2536 : บทคัดย่อ) ได้ศึกษาการเปรียบเทียบค่าสถิติของมาตรประเมินค่าที่มีจำนวนช่วงต่างกัน 5 แบบ คือ แบบ 7 ช่วง แบบ 6 ช่วง แบบ 5 ช่วง แบบ 4 ช่วง และแบบ 3 ช่วง ที่ได้จากการประเมินการสำรวจเจตคติในการเรียน กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนชั้นมัธยมศึกษา ปีที่ 3 สังกัดกรมสามัญศึกษา จังหวัดสระบุรี จำนวน 695 คน ซึ่งได้มาโดยวิธีสุ่มแบบแบ่งแล้วสุ่ม 2 ชั้น ออกเป็น 5 กลุ่มๆ ละ 139 คน ผลการวิจัยพบว่า ค่าเฉลี่ย ความแปรปรวนและค่าความเชื่อมั่นแบบวัดซ้ำของมาตรประเมินที่มีจำนวนมาตราต่างกัน 5 แบบ แตกต่างกัน

เกศเกล้า หมุ่มมิ่ง (2539 : บทคัดย่อ) ได้ศึกษาสูตรหาค่าความเชื่อมั่น จำนวน 6 สูตร ในการประมาณค่าความเชื่อมั่นของมาตรประเมินค่าด้านมโนภาพแห่งตน ซึ่งได้แก่สูตร α , r_{11} , r_{12} , α_{λ} , Ω และ Ω_w กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2538 ของโรงเรียนขององค์กรสังฆมณฑลกรุงเทพฯ ที่ตั้งอยู่ในกรุงเทพมหานคร จำนวน 600 คน โดยวิธีการสุ่มแบบแบ่งชั้น ผลการวิจัยพบว่า

1) ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นด้วยสูตร α มีค่าต่ำสุดและสูตร r_{L2} มีค่าสูงสุด

2) ความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นที่คำนวณด้วยสูตร r_{L1} มีค่าต่ำสุดและสูตร α มีค่าสูงสุด

3) การเปรียบเทียบค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตรทั้ง 6 สูตร จากกลุ่มตัวอย่าง 600 คน และกลุ่มตัวอย่างขนาด 40 คน จำนวน 200 กลุ่ม พบว่า ค่าความเชื่อมั่นที่คำนวณด้วยสูตร α มีค่าต่ำสุด และสูตร Ω_w มีค่าสูงสุด เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่น โดยใช้สูตร UX_1 พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จากนั้นทดสอบค่าความเชื่อมั่นรายคู่ด้วยวิธีการของมาราคูโยโล พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกคู่ ส่วนค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตรทั้ง 6 สูตร พบว่า ค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตร α มีค่าต่ำสุดและสูตร Ω_w มีค่าสูงสุด ทดสอบความแตกต่างของค่าความเชื่อมั่นเฉลี่ย โดยใช้สูตรการวิเคราะห์ความแปรปรวนแบบตัวแปรพหุทางเดียว พบว่า มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

มาลี เดชะไตรศร (2537 : บทคัดย่อ) ได้ศึกษาปฏิสัมพันธ์ของจำนวนช่วงกับประเภทของคำตอบตรงกลางที่มีต่อการเลือกตอบคำตอบตรงกลางในมาตราประเมินค่า เมื่อกำหนดจำนวนช่วงกับประเภทของคำตอบตรงกลางต่างกัน กลุ่มตัวอย่างเป็นครู – อาจารย์ในโรงเรียนประถมศึกษาสังกัดสำนักงานคณะกรรมการ การประถมศึกษาแห่งชาติ กรุงเทพมหานคร จำนวน 342 คน ใช้แบบสอบถามความคิดเห็นเกี่ยวกับภาพพจน์ของครูเป็นเครื่องมือในการเก็บรวบรวมข้อมูล ผลการวิจัยพบว่า

1. มีปฏิสัมพันธ์ของจำนวนช่วงกับประเภทของคำตอบตรงกลางต่อการเลือกตอบคำตอบตรงกลางในมาตราประเมินค่าอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. เมื่อจำนวนช่วงเป็น 3 ช่วง การเลือกตอบคำตอบตรงกลางของแบบ 3 ช่วง – เป็นกลางและ 3 ช่วง – ไม่แน่ใจสูงกว่าแบบ 3 ช่วง – ไม่มีความคิดเห็น อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. เมื่อจำนวนช่วงเป็น 5 ช่วง การเลือกตอบคำตอบตรงกลางของแบบ 5 ช่วง – เป็นกลางและ 5 ช่วง – ไม่แน่ใจสูงกว่าแบบ 5 ช่วง – ไม่มีความคิดเห็น อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4. เมื่อจำนวนช่วงเป็น 7 ช่วง การเลือกตอบคำตอบตรงกลางไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

5. เมื่อประเภทของคำตอบตรงกลางเป็นคำว่า “เป็นกลาง” การเลือกตอบคำตอบตรงกลางของแบบ 3 ช่วง – เป็นกลางสูงกว่าแบบ 5 ช่วง – เป็นกลาง และ 7 ช่วง – เป็นกลาง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

6. เมื่อประเภทของคำตรงกลางเป็นคำว่า "ไม่แน่ใจ" การเลือกตอบคำตรงกลางของแบบ 3 ช่วง - ไม่แน่ใจสูงกว่าแบบ 7 ช่วง - ไม่แน่ใจ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

7. เมื่อประเภทของคำตรงกลางที่เป็นคำว่า "ไม่มีความคิดเห็น" การเลือกตอบคำตรงกลาง ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ฉวีวรรณ ครองยุทธ. (2539 : บทคัดย่อ) ได้ศึกษาการประมาณค่าความเชื่อมั่นของมาตรประเมินค่าที่มีขนาดกลุ่มตัวอย่างต่างกัน กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2538 ในจังหวัดอุดรธานี จำนวน 1,015 คน โดยวิธีการสุ่มแบบแบ่งชั้นและใช้เป็นประชากรเทียม ผลการวิจัยพบว่า

1. การเปรียบเทียบค่าความเชื่อมั่นที่คำนวณด้วยสูตร α_λ , Ω , Ω_w จากกลุ่มประชากรเทียม 1,015 คน พบว่า ค่าความเชื่อมั่นที่คำนวณด้วยสูตร α_λ มีค่าต่ำที่สุด และค่าความเชื่อมั่นที่คำนวณด้วยสูตร Ω_w มีค่าสูงสุด เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นโดยใช้สูตร UX_1 พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 จากนั้นทดสอบค่าความเชื่อมั่นรายคู่ด้วยวิธีการของ มาราคุโยโล พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกคู่

2. ค่าความเชื่อมั่นเฉลี่ยที่คำนวณจากสูตรเดียวกันเมื่อกลุ่มตัวอย่างมีขนาดต่างกันคือ 40, 80, 160, 250, 500 คน ให้ค่าความเชื่อมั่นเฉลี่ยแตกต่างกันทั้ง 3 สูตร และพบว่าสูตร α_λ และสูตร Ω เหมาะสมกับกลุ่มตัวอย่างขนาด 500 คน และสูตร Ω_w เหมาะสมกับกลุ่มตัวอย่างขนาด 500 คน และสูตร Ω_w เหมาะสมกับกลุ่มตัวอย่างขนาด 80 คน โดยพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานและค่า ความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นที่ต่ำเมื่อเทียบกับกลุ่มอื่น

พนิดา พุ่งเกียรติินาสุข (2540 : บทคัดย่อ) ได้ศึกษาการประมาณค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบการตอบแตกต่างกัน 3 รูปแบบ ได้แก่ แบบตัวเลขแบบกราฟฟิค และแบบบรรยาย โดยใช้สูตรในการประมาณค่าความเชื่อมั่น 2 สูตร คือ r_{L2} และ Ω_w กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของกลุ่มโรงเรียนมัธยมศึกษาส่วนกลาง กลุ่ม 4 กรุงเทพมหานคร ผลการศึกษาพบว่า

1. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความเชื่อมั่นของมาตรประเมินค่าทั้ง 3 รูปแบบ ที่คำนวณด้วยสูตร r_{L2} และ Ω_w มีค่าตั้งแต่ 0456 ถึง .0617

2. ค่าความลำเอียงทางสถิติในการประมาณค่าความเชื่อมั่นของมาตรประเมินค่าทั้ง 3 รูปแบบ ที่คำนวณด้วยสูตร r_{L2} และ Ω_w มีค่าตั้งแต่ -.0023 ถึง .0426

3. มาตรฐานประเมินค่าแบบกราฟิกมีค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยสูงสุดเมื่อคำนวณจากทั้งสูตร r_{L2} และ Ω_w เมื่อเปรียบเทียบค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยของมาตรฐานประเมินค่าทั้ง 3 รูปแบบ เมื่อคำนวณด้วยสูตรเดียวกัน พบว่า ทุกรูปแบบแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($P < .01$, $P < .05$) แต่เมื่อคำนวณด้วยสูตร Ω_w พบว่า แบบตัวเลขกับแบบบรรยายมีค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. เมื่อเปรียบเทียบค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยของมาตรฐานประเมินค่าแต่ละรูปแบบที่คำนวณจากสูตรต่างกันคือ พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($P < .01$) และสูตร Ω_w ให้ค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยสูงกว่าสูตร r_{L2} ในมาตรวัดประเมินค่าทุกรูปแบบ

บวร คมสิริพิทักษ์ (2541 : บทคัดย่อ) ได้ศึกษาการเปรียบเทียบความคลาดเคลื่อนแบบปล่อยคะแนน ความคลาดเคลื่อนแบบฮาโล และความคลาดเคลื่อนแบบขัดแย้ง ของมาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐานต่างกัน เมื่อสถานภาพของผู้ประเมินต่างกัน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ดัดแปลงมาจากแบบประเมินคุณภาพการสอนของครู ของอมรรัตน์ อลัมภินวัณ (2523) มีลักษณะเป็นมาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐาน 4 ช่วง 5 ช่วง 6 ช่วง และ 7 ช่วง กลุ่มตัวอย่างที่ใช้เป็นครูหมวดวิทยาศาสตร์ จำนวน 13 คน และนักเรียนสายวิทย์ - คณิต ระดับชั้นมัธยมศึกษาปีที่ 5 จำนวน 27 คน โดยให้กลุ่มตัวอย่างทำการประเมินคุณภาพการสอนของครูวิทยาศาสตร์ 2 คน ผลการวิจัยพบว่า

1. กรณีสถานภาพของผู้ประเมินเป็นนักเรียน ทำการประเมินคุณภาพการสอนของครู พบว่า มาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐาน 4 ช่วง 6 ช่วง และ 7 ช่วง มีความคลาดเคลื่อนแบบปล่อยคะแนนในขณะที่มาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐาน 5 ช่วง ไม่พบความคลาดเคลื่อนทั้ง 3 ชนิด

2. กรณีสถานภาพของผู้ประเมินเป็นเพื่อนร่วมงาน ทำการประเมินคุณภาพการสอนของครู พบว่า มาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐาน 4 ช่วง และ 6 ช่วง มีความคลาดเคลื่อนแบบปล่อยคะแนนในขณะที่มาตรฐานประเมินค่าที่มีจำนวนช่วงมาตรฐาน 5 ช่วง และ 7 ช่วง ไม่พบความคลาดเคลื่อนทั้ง 3 ชนิด

3. ไม่พบปฏิสัมพันธ์ระหว่างจำนวนช่วงมาตรฐานกับ สถานภาพของผู้ประเมิน ในการเกิดความ คลาดเคลื่อนทั้ง 3 ชนิด

4.2 งานวิจัยต่างประเทศ

โนวิก และเลวิส (Novick and Lewis. 1967 : 1-13) ได้แสดงที่มาของสัมประสิทธิ์แอลฟา (α) โดยใช้วิธีการทางคณิตศาสตร์และได้แสดงให้เห็นว่าสัมประสิทธิ์แอลฟาต้องมีเงื่อนไขที่จำเป็นและเพียงพอ (Necessary and Sufficient) จึงจะทำให้สัมประสิทธิ์แอลฟา มีค่าเท่ากับค่าความเชื่อมั่นของแบบทดสอบ นั่นก็คือสูตรสัมประสิทธิ์แอลฟากับสูตรความเชื่อมั่นในทางคณิตศาสตร์แล้วไม่ใช่สูตรเดียวกัน (not mathematical identity) สัมประสิทธิ์แอลฟาจะเท่ากับค่าความเชื่อมั่นของแบบทดสอบที่ประกอบด้วยส่วนย่อยได้ก็ต่อเมื่อส่วนย่อยของแบบทดสอบ หรือข้อสอบแต่ละข้อเป็นไปตามข้อตกลงที่ว่า คะแนนจริงของแต่ละส่วนมีความสัมพันธ์เป็นเส้นตรง ซึ่งหมายความว่าจำเป็นต้องมีคะแนนจริงสมมูลกัน (Essentially Tau-Equivalent)

ฟินน์ (Finn . 1972 : 255-265) ได้ศึกษาเปรียบเทียบค่าเฉลี่ยและค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบคำตอบต่างกัน โดยแบ่งการศึกษออกเป็น 2 กรณี คือ กรณีที่หนึ่ง ใช้กลุ่มตัวอย่างเป็นนิสิตระดับปริญญาตรีจำนวน 30 คน แบ่งเข้ากลุ่ม กลุ่มละ 10 คน เพื่อประเมินงาน 10 ชิ้น ซึ่งเครื่องมือเป็นมาตรประเมินค่าจำนวน 5 สเกล ที่มีรูปแบบต่างกัน 3 รูปแบบ คือ แบบที่มีคำอธิบายทุกสเกล แบบที่มี คำอธิบายเฉพาะสเกลที่ 2 กับ 4 และแบบที่ไม่มีคำอธิบายบนสเกลนอกจากตัวเลข ผลปรากฏว่าค่าเฉลี่ยของแต่ละงานแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่ค่าเฉลี่ยของแต่ละรูปแบบแตกต่างกันอย่างไม่มีนัยสำคัญ และเมื่อเปรียบเทียบค่าความเชื่อมั่น ผลปรากฏว่าแตกต่างกันอย่างไม่มีนัยสำคัญ กรณีที่สอง กลุ่มตัวอย่างคือนิสิตปริญญาตรี จำนวน 20 คน สุ่มเข้ากลุ่ม 4 กลุ่ม กลุ่มละ 5 คน เพื่อประเมินงาน 4 ชิ้น โดยมีรูปแบบ คำตอบของมาตรประเมินค่าแตกต่างกัน 4 รูปแบบคือ แบบไม่มีคำอธิบาย แบบมีคำอธิบายสเกล ตรงกลาง แบบที่มีคำอธิบายสเกลที่ 1 กับ 3 และแบบที่มีคำอธิบายสเกลที่ 1 กับ 5 ในแต่ละรูปแบบจะมีจำนวนสเกลต่างกันไป 4 แบบคือ 3, 5, 7, และ 9 สเกล ตามลำดับ ผลปรากฏว่าทั้งค่าเฉลี่ยและค่าความเชื่อมั่นในประเภทของงาน และจำนวนช่วงแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่รูปแบบคำตอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ไอเคน (Aiken. 1983 : Abstract) ได้ทำการศึกษาความสัมพันธ์ระหว่างจำนวนคำตอบกับความแปรปรวนของมาตรประเมินค่า ที่ใช้ประเมินการสอนของครู 10 ข้อ ซึ่งมีรูปแบบคำตอบแบบกราฟิก 5 ช่วง ที่เป็นแบบมาตรฐานโดยที่ 1 หมายถึง น้อยที่สุด และ 5 หมายถึง มากที่สุด ผู้วิจัยได้เปลี่ยนรูปแบบคำตอบเป็น 2 ช่วง, 3 ช่วง, 4 ช่วง, 5 ช่วง, 6 ช่วง และ 7 ช่วง โดยให้นักเรียนระดับต่ำกว่าปริญญา จำนวน 624 คน ประเมินการสอนของครู ผลการวิจัยพบว่า ค่าเฉลี่ยและค่าความแปรปรวนจะเพิ่มขึ้น เมื่อจำนวนช่วงของคำตอบของมาตรประเมินค่ามีมากขึ้น

นิวสเต็ด และอาร์โนล (Newstead and Arnold. 1989 : 33-43) ได้ศึกษาผลของรูปแบบคำตอบของมาตรประเมินค่าที่มีต่อผลการประเมินการสอนของครู โดยให้นักศึกษาชั้นปีที่ 2 จำนวน 44 คน และ นักศึกษาชั้นปีที่ 1 จำนวน 108 คน ประเมินประสิทธิภาพของผู้สอนโดยใช้แบบประเมินซึ่งเป็นมาตรประเมินค่า 3 รูปแบบ คือ แบบบรรยายระดับพฤติกรรมแบบกราฟิกที่มีคำอธิบายหัวท้าย และแบบตัวเลขที่เป็นเปอร์เซ็นต์ของพฤติกรรม ผลปรากฏว่าทั้ง 3 รูปแบบมีค่าเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

เบคอน ซาวเออร์ และ ยัง (Bacon, Sauer and Young. 1995 : 394-406) ได้ศึกษาเกี่ยวกับการประมาณค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์แอลฟา (α) สูตรโอเมกาการณีนีที่มีน้ำหนักองค์ประกอบเท่ากัน (Ω) และใช้สูตรโอเมกาการณีนีที่มีน้ำหนักองค์ประกอบไม่เท่ากัน (Ω_w) ทั้งสามสูตรใช้ค่าน้ำหนักองค์ประกอบ (Factor Loading) ในการประมาณค่าความเชื่อมั่น โดยมีวิธีการศึกษาดังนี้คือ ทำการวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อตรวจสอบความเป็นมิติเดียว (Unidimension) และกำหนดค่าน้ำหนักองค์ประกอบ ข้อคำถามที่ใช้ในการศึกษาครั้งนี้ จะต้องมียาค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .3 - .9 และจำนวนข้อคำถาม ที่ใช้จะอยู่ระหว่าง 3 - 15 ข้อ ผลจากการศึกษาพบว่า ถ้ากำหนดชุดของน้ำหนักองค์ประกอบให้ 2 ชุด ดังนี้ .95, .95, .95, .01 และ .95, .95, .95 แล้วใช้สูตรโอเมกา (Ω_w) ในการประมาณค่าความเชื่อมั่น ข้อมูลชุดแรกจะได้ .9652408 ขณะที่ชุดที่สอง จะได้ .9652406 แม้ว่าจะเกิดความแตกต่างเพียงเล็กน้อยแต่ก็มีผลต่อค่าความเชื่อมั่นที่ลดลง ซึ่งเป็นผลเนื่องมาจากการตัดข้อคำถามที่มีค่าน้ำหนักองค์ประกอบน้อยเกินไป ดังนั้นการใช้สูตรแอลฟา (α) หรือสูตรโอเมกา (Ω) จะทำให้เกิดความคลาดเคลื่อนที่คล้ายคลึงกับความ คลาดเคลื่อนแบบที่ 1 (Type I error) คือ ปฏิเสธข้อคำถามที่ควรยอมรับได้ ผลจากการวิจัยอีกประการหนึ่งก็คือ ในกรณีที่จำนวนข้อคำถามน้อยและมีช่วงของค่าความเชื่อมั่นกว้างแล้ว ตัวอย่างเช่น ข้อคำถาม 3 ข้อ มีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .3 - .9 สูตรแอลฟา (α) และสูตรโอเมกา (Ω) จะให้ค่าความเชื่อมั่นเท่ากับ .596 และ .651 ตามลำดับ ขณะที่สูตรโอเมกา (Ω_w) จะให้ค่าความเชื่อมั่นเท่ากับ .831 ดังนั้นการใช้สูตรโอเมกา (Ω_w) ในการประมาณค่าความเชื่อมั่นจึงมีความเหมาะสมมากที่สุด

ซิมเมอร์แมน ซัมโบ และลาลอนด์ (Zimmerman, Zumbo and Lalond. 1993 : 33-49) ได้ศึกษาคำแนะนำโน้มเข้าส่วนกลางและการกระจายของสัมประสิทธิ์แอลฟาที่คำนวณจากคะแนนซึ่งจำลองด้วยคอมพิวเตอร์ (Computer Simulation) ให้ละเมิดข้อตกลงของสูตรสัมประสิทธิ์แอลฟา 2 ข้อ คือ 1) คะแนนที่ได้ไม่เป็นไปตามข้อตกลงของคะแนนจริงสมมูล และ 2) คะแนนคลาดเคลื่อนไม่เป็นไปตามข้อตกลงที่ว่าต้องมีความอิสระกัน ผลการวิจัยพบว่า เมื่อคะแนนเป็นไปตามข้อตกลงแล้วค่าเฉลี่ยของสัมประสิทธิ์แอลฟามีค่าใกล้เคียงกับสัมประสิทธิ์ที่คำนวณจากประชากร แต่หาที่ได้มีการกระจายสูงมาก สัมประสิทธิ์แอลฟาจะมีค่าต่ำกว่าค่าความเชื่อมั่น (Underestimated Reliability) เมื่อละเมิดข้อตกลงเกี่ยวกับคะแนนจริงสมมูล และ

สัมประสิทธิ์ แอลฟาจะมีค่าสูงกว่าค่าความเชื่อมั่น เมื่อละเมิดข้อตกลงเกี่ยวกับคะแนนคลาดเคลื่อน ต้องอิสระจากกัน และทั้งสองกรณีมีความลำเอียงของการประมาณค่า มีความผันแปรขึ้นอย่างเป็นระบบตามระดับของการละเมิดข้อตกลง ข้อค้นพบนี้สะท้อนให้เห็นในทางปฏิบัติว่า สัมประสิทธิ์แอลฟา ที่ปรากฏอยู่ในหนังสือคู่มือดำเนินการสอบ รายงานการวิจัยและที่อื่นๆ บางครั้งอาจมีค่าสูงกว่าและบางครั้งอาจมีค่าต่ำกว่าสัมประสิทธิ์ของความเชื่อมั่นของประชากร (บุญเชิด ภิญโญนนันทพงษ์. 2538 : 57 – 58 ; อ้างอิงจาก Zimmerman, Zumbo and Lalonde. 1993 : *Education and Psychological Measurement.*)

จากการศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง สรุปได้ว่ามาตรประเมินค่า มีได้หลายรูปแบบ ในการสร้างนั้นผู้สร้างควรมีความระมัดระวังเพื่อช่วยลดความคลาดเคลื่อน ที่อาจเกิดขึ้นได้ โดยในส่วนของลักษณะที่จะให้ประเมินต้องมีความชัดเจนผู้ตอบเข้าใจความหมาย ไปในทิศทางเดียวกัน และในส่วนของคำตอบก็ต้องกำหนดรูปแบบคำตอบและระดับในการ ประเมินไว้ให้ชัดเจน นอกจากนี้ยังพบว่ามีการนำมาตรประเมินค่าไปใช้มากในการวัดคุณลักษณะ ทางจิตวิทยาหรือบุคลิกภาพทางสังคม จึงมีผู้พยายามที่จะปรับปรุงพัฒนามาตรประเมินค่าเพื่อ ให้เกิดประสิทธิภาพมากที่สุดในการนำไปใช้ทั้งในและต่างประเทศ และสิ่งที่พวกเขาสนใจศึกษา กันมากคือการพัฒนาในส่วนของคำตอบ ทำให้เกิดรูปแบบคำตอบมากมาย โดยพบว่าส่วนใหญ่ นิยมประมาณค่าความเชื่อมั่นของมาตรประเมินค่าโดยใช้สัมประสิทธิ์แอลฟา (α) ของครอนบัค (Cronbach) ในความเป็นจริงแล้วสัมประสิทธิ์แอลฟาที่คำนวณได้จะเท่ากับค่าความเชื่อมั่นของ แบบทดสอบที่ประกอบด้วยส่วนย่อยได้ก็ต่อเมื่อส่วนย่อยของแบบทดสอบเป็นไปตามข้อตกลงของ คะแนนจริงสมมูล (Essentially Tau Equivalent) แต่ในทางปฏิบัติไม่อาจจะทำได้ ดังนั้นการใช้ สัมประสิทธิ์แอลฟาจึงเป็นการละเมิดข้อตกลง ดังกล่าว แล้วจะส่งผลทำให้สัมประสิทธิ์แอลฟามี ค่าต่ำกว่าหรือสูงกว่าสัมประสิทธิ์ความเชื่อมั่นของประชากร ดังนั้น ในการประมาณความเชื่อมั่น ของมาตรประเมินค่าควรพิจารณาให้รอบคอบ ซึ่งการประมาณค่าความเชื่อมั่นโดยใช้ข้อตกลง ตามแบบจำลองคะแนนจริงสัมพันธ์จะเป็นวิธีการที่เหมาะสม เป็นวิธีที่มีความปลอดภัยหรือมี ความเสี่ยงน้อยที่สุดที่จะให้ผลคลาดเคลื่อน ซึ่งจากการค้นคว่ายังไม่พบว่ามีผู้ศึกษามาตรประเมินค่า รูปแบบคำตอบใดที่มีจำนวนมาตราเท่าไรจึงจะมีความเชื่อมั่นมากที่สุด จากผลการศึกษาที่กล่าวมา ทั้งหมดจะเห็นว่ามาตรประเมินค่าที่มีรูปแบบคำตอบแตกต่างกันหรือที่มีจำนวนมาตราต่างกัน จะมีผลต่อคำตอบที่ได้ ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงสนใจที่จะศึกษาความเชื่อมั่นของ มาตรประเมินค่าที่มีรูปแบบคำตอบและจำนวนมาตราต่างกัน เพื่อจะศึกษาว่าในรูปแบบคำตอบใด ควรจะมีมาตรากี่มาตราจึงจะมีความเชื่อมั่นมากที่สุด โดยใช้สูตรสัมประสิทธิ์แอลฟา (α) ของครอนบัค (Cronbach) และสูตรของเฟลด์ ราซ

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้ มีความมุ่งหมายเพื่อศึกษาว่ามาตรฐานประเมินค่าที่มีรูปแบบมาตรฐานประเมินค่า 3 รูปแบบ ได้แก่ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก รูปแบบใดมีค่าความเชื่อมั่นที่มากที่สุด และจะมีค่าความเชื่อมั่นแตกต่างกันหรือไม่เมื่อมีจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มีรายละเอียดและขั้นตอนการดำเนินงานวิจัยดังนี้

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 113 โรงเรียน จำนวนห้องเรียน 1,204 ห้อง จำนวนนักเรียน 49,836 คน

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ของโรงเรียนมัธยมศึกษา ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา ซึ่งได้มาโดยวิธีสุ่มหลายขั้นตอน (Multistage Random Sampling) และประมาณกลุ่มตัวอย่างด้วยความเชื่อมั่น 95% ที่ความคลาดเคลื่อน 10% เมื่อเปรียบเทียบตารางกำหนดขนาดของกลุ่มตัวอย่างสำหรับศึกษาค่าเฉลี่ยประชากรพบว่าต้องใช้กลุ่มตัวอย่างจำนวน 397 คน (ศิริชัย กาญจนวาสี และคณะ. 2537 : 106) เพื่อความเหมาะสมผู้วิจัยจึงเพิ่มจำนวนกลุ่มตัวอย่างเป็น 540 คน และเพื่อให้ข้อมูลมีการกระจายมากเกินไปเนื่องจากต่างโรงเรียน จึงแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม กลุ่มละ 3 โรงเรียน โรงเรียนละ 3 ห้องเรียน ดังนี้ กลุ่มที่ 1 คือ กลุ่มสำหรับตอบมาตรฐานประเมินค่าที่มีจำนวนมาตราเป็นจำนวนคู่ คือ 4, 6 และ 8 มาตรา โดยมีรูปแบบแบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก กลุ่มที่ 2 กลุ่มสำหรับตอบมาตรฐานประเมินค่าที่มีจำนวนมาตราเป็นจำนวนคี่ คือ 5, 7 และ 9 มาตรา โดยมีรูปแบบแบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก การสุ่มตัวอย่างใช้วิธีสุ่มหลายขั้นตอน (Multistage Random Sampling) ดังนี้

ขั้นที่ 1 สุ่มเลือกโรงเรียนจำนวน 6 โรงเรียน โดยวิธีการสุ่มตัวอย่างง่าย

ขั้นที่ 2 แล้วทำการสุ่มโรงเรียนเข้ากลุ่มเครื่องมือจากทั้งหมด 2 กลุ่ม คือมาตรฐานประเมินค่าที่มีจำนวนมาตราเป็นจำนวนคู่และจำนวนคี่ กลุ่มละ 3 โรงเรียน โดยวิธีการสุ่มอย่างง่าย

ขั้นที่ 3 สุ่มรูปแบบมาตรฐานประเมินค่าคือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จากกลุ่มที่ 1 ให้แต่ละโรงเรียน โดยวิธีการสุ่มอย่างง่าย

ขั้นที่ 4 สุ่มรูปแบบมาตรฐานประเมินค่าคือ แบบเดิมตัวเลข แบบบรรยาย แบบกราฟฟิก จากกลุ่มที่ 2 ให้แต่ละโรงเรียน โดยวิธีการสุ่มอย่างง่าย

ขั้นที่ 5 สุ่มห้องเรียนจากโรงเรียนกลุ่มตัวอย่าง โรงเรียนละ 3 ห้องเรียน โดยวิธีการสุ่มอย่างง่าย

ขั้นที่ 6 สุ่มนักเรียน ห้องเรียนละ 30 คน โดยวิธีการสุ่มอย่างง่าย ทำให้ได้จำนวนนักเรียนโรงเรียนละ 90 คน รวมจำนวนกลุ่มตัวอย่างทั้งหมด 540 คน รายละเอียดปรากฏดังแสดงในตาราง 1

ตาราง 1 จำนวนนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนมัธยมศึกษาในกรุงเทพมหานคร

ชื่อโรงเรียน	จำนวนห้องเรียน	จำนวนนักเรียน ในแต่ละห้อง	จำนวนนักเรียน ทั้งหมด
1. โรงเรียนปทุมคงคา	3	30	90
2. โรงเรียนสายน้ำผึ้ง	3	30	90
3. โรงเรียนวัดธาตุทอง	3	30	90
4. โรงเรียนหอวัง	3	30	90
5. โรงเรียนพระราชราษฎร์อุปถัมภ์	3	30	90
6. โรงเรียนกุนนทีรุชาราภวิทยา	3	30	90

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นมาตรฐานประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบ ของจอห์น และ โรดวอลท์ (Jones and Rhodewalt. 1982) จำนวน 25 ข้อ เป็นข้อทางบวก 17 ข้อข้อทางลบ 8 ข้อ

ขั้นตอนการพัฒนาเครื่องมือ

1. ผู้วิจัยนำมาตรฐานประเมินค่าของลักษณะนิสัยการทำให้ตนเองเสียเปรียบของจอห์น และโรดวอลท์ (Jones and Rhodewalt. 1982) จำนวน 25 ข้อ มาแปลและเรียบเรียงเป็นภาษาไทย

2. นำมาตรฐานประเมินค่าจากข้อ 1 ไปให้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ตรวจสอบความเที่ยงตรงเชิงพิณิจ (Face Validity) ในการพิจารณาถ้าผู้เชี่ยวชาญมีความเห็นสอดคล้องกันเป็นรายข้อความ 3 ใน 5 ท่าน ถือว่าข้อความนั้นใช้ได้ ซึ่งผลจากผู้เชี่ยวชาญพิจารณาได้ผลดังนี้

2.1 ผู้เชี่ยวชาญเสนอแนะให้ผู้วิจัยปรับปรุงข้อคำถามที่ถามหลายทิศทาง คือ ข้อ 3 และข้อ 8 ให้เป็นข้อคำถามที่ถามในทิศทางเดียว ซึ่งผู้วิจัยได้ปรับปรุงตามที่เสนอแนะปรากฏว่าได้ข้อคำถามทั้งหมด 27 ข้อ

2.2 มาตรฐานประเมินค่า จำนวน 27 ข้อ ผ่านการคัดเลือกทั้ง 27 ข้อ เป็นข้อความทางบวก 19 ข้อ ทางลบ 8 ข้อ โดยมีค่า IOC ระหว่าง .60 ถึง 1.00

3. นำมาตรฐานประเมินค่าของลักษณะนิสัยการทำให้ตนเองเสียเปรียบของจอห์นและโรดวอลท์ (Jones and Rhodewalt. 1982) ที่ผู้วิจัยแปลเรียบเรียง ปรับปรุงและตรวจสอบความเที่ยงตรงแล้ว ไปทดสอบครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนเทพลีลาที่ไม่ใช่กลุ่มตัวอย่างจำนวน 200 คน แล้วนำมาตรวจให้คะแนน จากนั้นจึงนำผลที่ได้มาวิเคราะห์หาค่าอำนาจจำแนกรายข้อโดยใช้สัมประสิทธิ์สหสัมพันธ์ของข้อคำถามกับคะแนนรวมจากข้ออื่น ๆ ที่เหลือทั้งหมด โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Product-Moment Correlation) r_{xy} เพื่อคัดเลือกข้อที่มีค่า r_{xy} ตั้งแต่ 0.20 ขึ้นไป ที่ระดับนัยสำคัญ .01 ผลปรากฏว่ามาตรฐานประเมินค่าทั้ง 27 ข้อ มีค่าอำนาจจำแนกอยู่ระหว่าง .991-.997

4. นำมาตรฐานประเมินค่าของลักษณะนิสัยการทำให้ตนเองเสียเปรียบของจอห์นและโรดวอลท์ (Jones and Rhodewalt. 1982) ที่คัดเลือกในการทดสอบครั้งที่ 1 ไปทดสอบครั้งที่ 2 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนมักกะสัน ที่ไม่ใช่กลุ่มตัวอย่างจำนวน 200 คน แล้วนำมาตรวจให้คะแนนเพื่อหาค่าความเชื่อมั่นของมาตรฐานประเมินค่าโดยวิธีหาค่าสัมประสิทธิ์แอลฟา (α -Coefficient) ซึ่งปรากฏว่าได้ค่าความเชื่อมั่นของมาตรฐานประเมินค่า 0.7738

5. นำมาตรฐานประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบที่ได้จากการทดสอบครั้งที่ 2 มาจัดเป็นชุดของมาตรวัด โดยดัดแปลงรูปแบบการตอบของมาตรวัดเป็น 3 รูปแบบ คือ มาตรฐานประเมินค่าแบบแบบเต็มตัวเลข แบบกราฟฟิก และแบบบรรยาย โดยทั้ง 3 รูปแบบมีจำนวน 4, 5, 6, 7, 8 และ 9 มาตรวัด ดังตัวอย่างต่อไปนี้

ฉบับที่ 1 แบบเต็มตัวเลข 4 มาตรวัด

โปรดเติมตัวเลข 1-4 ในช่องว่างหน้าแต่ละข้อความ ตามที่ท่านเห็นว่าข้อความนั้นๆ บรรยายถึงลักษณะซึ่งตรงกับตัวท่านเองมากที่สุด

- | | |
|----------------------|-------------------------|
| ระดับมาตรา 4 มาตรวัด | 1. ไม่เห็นด้วยมากที่สุด |
| | 2. ไม่เห็นด้วย |
| | 3. เห็นด้วย |
| | 4. เห็นด้วยมากที่สุด |

_____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก

_____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 2 แบบเดิมตัวเลข 5 มาตรา

โปรดเติมตัวเลข 1 – 5 ในช่องว่างหน้า แต่ละข้อความ ตามที่ท่านเห็นว่าข้อความนั้นๆ บรรยายถึงลักษณะซึ่งตรงกับตัวท่านเองมากที่สุด

- | | |
|--------------------|---|
| ระดับมาตรา 5 มาตรา | 1. ไม่เห็นด้วยมากที่สุด
2. ไม่เห็นด้วย
3. ไม่แน่ใจ
4. เห็นด้วย
5. เห็นด้วยมากที่สุด |
|--------------------|---|

- ____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก
- ____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 3 แบบเดิมตัวเลข 6 มาตรา

โปรดเติมตัวเลข 1-6 ในช่องว่างหน้าแต่ละข้อความ ตามที่ท่านเห็นว่าข้อความนั้นๆ บรรยายถึงลักษณะที่ตรงกับตัวท่านเองมากที่สุด

- | | |
|--------------------|---|
| ระดับมาตรา 6 มาตรา | 1. ไม่เห็นด้วยมากที่สุด
2. ไม่เห็นด้วย
3. ค่อนข้างไม่เห็นด้วย
4. ค่อนข้างเห็นด้วย
5. เห็นด้วย
6. เห็นด้วยมากที่สุด |
|--------------------|---|

- ____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก
- ____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 4 แบบเดิมตัวเลข 7 มาตรา

โปรดเติมตัวเลข 1 – 7 ในช่องว่างหน้าแต่ละข้อความตามที่ท่านเห็นว่าข้อความนั้นๆ บรรยายถึงลักษณะซึ่งตรงกับตัวท่านเองมากที่สุด

- | | |
|-------------------------|--|
| จำนวนระดับมาตรา 7 มาตรา | 1. ไม่เห็นด้วยมากที่สุด
2. ไม่เห็นด้วย
3. ค่อนข้างไม่เห็นด้วย
4. ไม่แน่ใจ
5. ค่อนข้างเห็นด้วย
6. เห็นด้วย
7. เห็นด้วยมากที่สุด |
|-------------------------|--|

- _____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก
 _____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 5 แบบเดิมตัวเลข 8 มาตรา

โปรดเติมตัวเลข 1 – 8 ในช่องว่างหน้าแต่ละข้อความตามที่ท่านเห็นว่าข้อความนั้น uly
 บรรยายถึงลักษณะซึ่งตรงกับตัวท่านเองมากที่สุด

จำนวนระดับมาตรา 8 มาตรา

1. ไม่เห็นด้วยมากที่สุด
2. ไม่เห็นด้วยค่อนข้างมาก
3. ไม่เห็นด้วยค่อนข้างน้อย
4. ไม่เห็นด้วยน้อยที่สุด
5. เห็นด้วยน้อยที่สุด
6. เห็นด้วยค่อนข้างน้อย
7. เห็นด้วยค่อนข้างมาก
8. เห็นด้วยมากที่สุด

- _____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก
 _____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 6 แบบเดิมตัวเลขของลักษณะนิสัยการทำให้ตนเองเสียเปรียบของจอห์นและโรดวอลท์
 (Jones and Rhodewalt, 1982) 9 มาตรา

โปรดเติมตัวเลข 1-9 ในช่องว่างหน้าแต่ละข้อความตามที่ท่านเห็นว่าข้อความนั้น uly
 บรรยายถึงลักษณะซึ่งตรงกับตัวท่านเองมากที่สุด

จำนวนระดับมาตรา 9 มาตรา

1. ไม่เห็นด้วยมากที่สุด
2. ไม่เห็นด้วยค่อนข้างมาก
3. ไม่เห็นด้วยค่อนข้างน้อย
4. ไม่เห็นด้วยน้อยที่สุด
5. ไม่แน่ใจ
6. เห็นด้วยน้อยที่สุด
7. เห็นด้วยค่อนข้างน้อย
8. เห็นด้วยค่อนข้างมาก
9. เห็นด้วยมากที่สุด

- _____ 1. เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก
 _____ 2. ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ฉบับที่ 7 แบบบรรยาย 4 มาตรา

ให้นักเรียนพิจารณาข้อความในแต่ละข้อ แล้ว ✕ ลงในช่องใดก็ได้ที่ตรงกับความเป็นจริงมากที่สุด

ข้อความ	ไม่เห็นด้วยที่สุด	ไม่เห็นด้วย	เห็นด้วย	เห็นด้วยที่สุด
(0) เวลาที่ฉันทำอะไรผิดฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก (00) ถ้าฉันพยายามมากขึ้นฉันคงทำได้ดีกว่านี้				

ฉบับที่ 8 แบบบรรยาย 5 มาตรา

ข้อความ	ไม่เห็นด้วยที่สุด	ไม่เห็นด้วย	ไม่แน่ใจ	เห็นด้วย	เห็นด้วยที่สุด
(0) เวลาที่ฉันทำอะไรผิดฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก (00) ถ้าฉันพยายามมากขึ้นฉันคงทำได้ดีกว่านี้					

ฉบับที่ 9 แบบบรรยาย 6 มาตรา

ข้อความ	ไม่เห็นด้วยที่สุด	ไม่เห็นด้วย	ค่อนข้างไม่เห็นด้วย	ค่อนข้างเห็นด้วย	เห็นด้วย	เห็นด้วยที่สุด
(0) เวลาที่ฉันทำอะไรผิดฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก (00) ถ้าฉันพยายามมากขึ้นฉันคงทำได้ดีกว่านี้						

ฉบับที่ 10 แบบบรรยาย 7 มาตรา

ข้อความ	ไม่เห็น ด้วย ที่สุด	ไม่ เห็น ด้วย	ค่อนข้าง ไม่เห็น ด้วย	ไม่ แน่ใจ	ค่อนข้าง เห็นด้วย	เห็น ด้วย	เห็น ด้วย ที่สุด
(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพ แวดล้อมเป็นอันดับแรก (00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้							

ฉบับที่ 11 แบบบรรยาย 8 มาตรา

ข้อความ	ไม่ เห็น ด้วย ที่สุด	ไม่เห็น ด้วย ค่อนข้าง มาก	ไม่เห็น ด้วย ค่อนข้าง น้อย	ไม่เห็น ด้วย น้อยที่ สุด	เห็น ด้วย น้อย ที่สุด	เห็น ด้วย ค่อนข้าง น้อย	เห็น ด้วย ค่อนข้าง มาก	เห็น ด้วย มากที่สุด
(0) เวลาที่ฉันทำอะไร ผิดฉันมักจะโทษ สภาพแวดล้อม เป็นอันดับแรก (00) ถ้าฉันพยายาม มากขึ้นฉันคงทำ ได้ดีกว่านี้								

ฉบับที่ 12 แบบบรรยาย 9 มาตรา

ข้อความ	ไม่เห็นด้วยที่สุด	ไม่เห็นด้วย ค่อนข้างมาก	ไม่เห็นด้วย ค่อนข้างน้อย	ไม่เห็นด้วย น้อยที่สุด	ไม่แน่ใจ	เห็นด้วย น้อยที่สุด	เห็นด้วย ค่อนข้างน้อย	เห็นด้วย ค่อนข้างมาก	เห็นด้วย มากที่สุด
(0) เวลาที่ฉันทำอะไรผิดฉันมักจะโทษสภาพแวดล้อมเป็นอันดับแรก									
(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้									

ฉบับที่ 13 แบบกราฟฟิก 4 มาตรา

ให้นักเรียนพิจารณาข้อความในแต่ละข้อ แล้วทำเครื่องหมาย ✓ ลงในช่องใดก็ได้ที่ตรงกับความเป็นจริงมากที่สุด

(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

ไม่เห็นด้วยมากที่สุด

--	--	--	--

 เห็นด้วยมากที่สุด

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ไม่เห็นด้วยมากที่สุด

--	--	--	--

 เห็นด้วยมากที่สุด

ฉบับที่ 14 แบบกราฟฟิก 5 มาตรา

(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

ไม่เห็นด้วยมากที่สุด

--	--	--	--	--

 เห็นด้วยมากที่สุด

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ไม่เห็นด้วยมากที่สุด

--	--	--	--	--

 เห็นด้วยมากที่สุด

ฉบับที่ 15 แบบกราฟฟิก 6 มาตรา

(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

[illegible]

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

[illegible]

ฉบับที่ 16 แบบกราฟฟิก 7 มาตรา

(๐) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

[illegible]

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ไม่เห็นด้วยมากที่สุด | | | | | | | | | | เห็นด้วยมากที่สุด

ฉบับที่ 17 แบบกราฟฟิก 8 มาตรา

(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

ไม่เห็นด้วยมากที่สุด | | | | | | | | | | เห็นด้วยมากที่สุด

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

[illegible]

ฉบับที่ 18 แบบกราฟฟิก 9 มาตรา

(0) เวลาที่ฉันทำอะไรผิด ฉันมักจะโทษสภาพแวดล้อม เป็นอันดับแรก

ไม่เห็นด้วยมากที่สุด

--	--	--	--	--	--	--	--	--	--

 เห็นด้วยมากที่สุด

(00) ถ้าฉันพยายามมากขึ้น ฉันคงทำได้ดีกว่านี้

ไม่เห็นด้วยมากที่สุด

--	--	--	--	--	--	--	--	--	--

 เห็นด้วยมากที่สุด

การเก็บรวบรวมข้อมูลมีขั้นตอนดังนี้

ผู้วิจัยนำเครื่องมือการวิจัยทั้ง 18 ฉบับ ไปเก็บรวบรวมข้อมูลกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา ที่เป็นกลุ่มตัวอย่าง โดยได้ดำเนินการเก็บรวบรวมข้อมูลดังนี้

1. ทำหนังสือราชการขอความร่วมมือในการทำวิจัยจากบัณฑิตวิทยาลัย ถึงผู้บริหารสถานศึกษาของโรงเรียนในเขตกรุงเทพมหานครที่เป็นกลุ่มตัวอย่าง

2. ผู้วิจัยนำส่งจดหมายราชการจากบัณฑิตวิทยาลัยพร้อมนำตัวอย่างมาตรประเมินค่าที่จะใช้ในการดำเนินการเก็บรวบรวมข้อมูลในโรงเรียนนั้นถึงผู้บริหารสถานศึกษา เพื่อขอความร่วมมือในการเก็บรวบรวมข้อมูล พร้อมทั้งนัดวันและเวลาที่จะไปดำเนินการเก็บรวบรวมข้อมูล

3. ผู้วิจัยดำเนินการเก็บข้อมูลด้วยตนเอง โดยก่อนแจกแบบสอบถามได้ชี้แจงให้กลุ่มตัวอย่างนักเรียนเข้าใจว่าต้องการวัตถุประสงค์สภาพการทำให้ตนเองเสียเปรียบ เมื่อได้รับแบบสอบถามขอให้อ่านคำชี้แจงในการตอบให้เข้าใจ แล้วจึงตอบการแจกแบบสอบถาม ทั้งนี้ได้กำหนดให้ผู้ตอบไม่ต้องลงชื่อในแบบสอบถาม ซึ่งในการเก็บรวบรวมข้อมูลผู้วิจัยได้ดำเนินการเก็บดังนี้

3.1 สุ่มเลือกโรงเรียนจำนวน 6 โรงเรียน แล้วทำการสุ่มโรงเรียนเข้ากลุ่มเครื่องมือจากทั้งหมดเป็น 2 กลุ่ม กลุ่มละ 3 โรงเรียน จากนั้นสุ่มรูปแบบมาตรประเมินค่าคือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จากกลุ่มที่ 1 และกลุ่มที่ 2 ให้แต่ละโรงเรียน โดยการวิธีสุ่มอย่างง่ายได้ดังนี้ คือ

กลุ่มที่ 1 สำหรับตอบมาตรฐานประเมินค่าที่มีมาตราเป็นจำนวนคู่ ได้แก่

-โรงเรียนที่ 1 คือ โรงเรียนมัธยมธาตุทองตอบมาตรฐานประเมินค่าแบบเต็มตัวเลข

-โรงเรียนที่ 2 คือ โรงเรียนสายน้ำผึ้งตอบมาตรฐานประเมินค่าแบบบรรยาย

-โรงเรียนที่ 3 คือ โรงเรียนกุนนทีรุชราภามวิทยาตอบมาตรฐานประเมินค่าแบบกราฟฟิก

กลุ่มที่ 2 สำหรับตอบมาตรฐานประเมินค่าที่มีมาตราเป็นจำนวนคี่ ได้แก่

-โรงเรียนที่ 4 คือ โรงเรียนพระราชราษฎร์อุปถัมภ์ตอบมาตรฐานประเมินค่า

แบบเต็มตัวเลข

-โรงเรียนที่ 5 คือ โรงเรียนปทุมคงคาตอบมาตรฐานประเมินค่าแบบบรรยาย

-โรงเรียนที่ 6 คือ โรงเรียนห่อวังตอบมาตรฐานประเมินค่าแบบกราฟฟิก

3.2 สุ่มเลือกห้องเรียน โรงเรียนละ 3 ห้องเรียน โดยวิธีการสุ่มอย่างง่าย

แล้วทำการเก็บข้อมูลในโรงเรียนที่ 1 โดยเก็บข้อมูลในแต่ละห้อง ดังนี้

ห้องเรียนที่ 1 นักเรียนตอบมาตรฐานประเมินที่มีจำนวนมาตรา 4, 6 และ 8

มาตรา ตามลำดับ

ห้องเรียนที่ 2 นักเรียนตอบมาตรฐานประเมินที่มีจำนวนมาตรา 6, 8 และ 4

มาตรา ตามลำดับ

ห้องเรียนที่ 3 นักเรียนตอบมาตรฐานประเมินค่าที่มีจำนวนมาตรา 8, 4 และ 6

มาตรา ตามลำดับ

แล้วทำการเก็บข้อมูลในโรงเรียนที่ 2 และ โรงเรียน 3 ในทำนองเดียวกับโรง

เรียนที่ 1

จากนั้นทำการเก็บข้อมูลในโรงเรียนที่ 4 โดยเก็บข้อมูลในแต่ละห้อง ดังนี้

ห้องเรียนที่ 1 นักเรียนตอบมาตรฐานประเมินที่มีจำนวนมาตรา 5, 7 และ 9

มาตรา ตามลำดับ

ห้องเรียนที่ 1 นักเรียนตอบมาตรฐานประเมินที่มีจำนวนมาตรา 7, 9 และ 5

มาตรา ตามลำดับ

ห้องเรียนที่ 1 นักเรียนตอบมาตรฐานประเมินที่มีจำนวนมาตรา 9, 5 และ 7

มาตรา ตามลำดับ

แล้วทำการเก็บข้อมูลในโรงเรียนที่ 5 และ โรงเรียน 6 ในทำนองเดียวกับโรงเรียนที่ 4 ปรากฏผลดังแสดงในภาพประกอบที่ 1

กลุ่มที่ 1				กลุ่มที่ 2			
โรงเรียน	ห้องเรียนที่			โรงเรียน	ห้องเรียนที่		
	1	2	3		1	2	3
โรงเรียนที่ 1 ตอบ มา ต ร ประเมินค่า แบบเติมตัวเลข	4 มาตรา ↓ 6 มาตรา ↓ 8 มาตรา	6 มาตรา ↓ 8 มาตรา ↓ 4 มาตรา	8 มาตรา ↓ 4 มาตรา ↓ 6 มาตรา	โรงเรียนที่ 4 ตอบ มา ต ร ประเมินค่า แบบเติมตัวเลข	5 มาตรา ↓ 7 มาตรา ↓ 9 มาตรา	7 มาตรา ↓ 9 มาตรา ↓ 5 มาตรา	9 มาตรา ↓ 5 มาตรา ↓ 7 มาตรา
โรงเรียนที่ 2 ตอบ มา ต ร ประเมินค่า แบบบรรยาย	4 มาตรา ↓ 6 มาตรา ↓ 8 มาตรา	6 มาตรา ↓ 8 มาตรา ↓ 4 มาตรา	8 มาตรา ↓ 4 มาตรา ↓ 6 มาตรา	โรงเรียนที่ 5 ตอบ มา ต ร ประเมินค่า แบบบรรยาย	5 มาตรา ↓ 7 มาตรา ↓ 9 มาตรา	7 มาตรา ↓ 9 มาตรา ↓ 5 มาตรา	9 มาตรา ↓ 5 มาตรา ↓ 7 มาตรา
โรงเรียนที่ 3 ตอบ มา ต ร ประเมินค่า แบบกราฟฟิก	4 มาตรา ↓ 6 มาตรา ↓ 8 มาตรา	6 มาตรา ↓ 8 มาตรา ↓ 4 มาตรา	8 มาตรา ↓ 4 มาตรา ↓ 6 มาตรา	โรงเรียนที่ 2 ตอบ มา ต ร ประเมินค่า แบบกราฟฟิก	5 มาตรา ↓ 7 มาตรา ↓ 9 มาตรา	7 มาตรา ↓ 9 มาตรา ↓ 5 มาตรา	9 มาตรา ↓ 5 มาตรา ↓ 7 มาตรา

ภาพประกอบ 1 แบบแผนการจัดกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล

4. ตรวจสอบสมบูรณ์ของข้อมูลที่ได้จากนักเรียนในแต่ละห้องเรียน
5. สุ่มนักเรียน ห้องเรียนละ 30 คน โดยวิธีการสุ่มอย่างง่าย ทำให้ได้จำนวนนักเรียนโรงเรียนละ 90 คน รวมจำนวนกลุ่มตัวอย่างทั้งหมด 540 คน
4. นำมาตรวัดประเมินค่า ตรวจสอบให้คะแนน โดยกำหนดให้คะแนน ดังนี้
ฉบับที่ 1 แบบเติมตัวเลข 4 มาตรา

ข้อความทางบวก		ข้อความทางลบ	
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 1 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 2 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 3 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 4 คะแนน

ฉบับที่ 2 แบบเติมตัวเลข 5 มาตรา

		ข้อความทางบวก	ข้อความทางลบ
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 1 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 2 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 3 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 4 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 5 คะแนน

ฉบับที่ 3 แบบเติมตัวเลข 6 มาตรา

		ข้อความทางบวก	ข้อความทางลบ
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 1 คะแนน
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 2 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 3 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 4 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 5 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 6 คะแนน

ฉบับที่ 4 แบบเติมตัวเลข 7 มาตรา

		ข้อความทางบวก	ข้อความทางลบ
เลือก	7	ให้คะแนน 7 คะแนน	ให้คะแนน 1 คะแนน
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 2 คะแนน
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 3 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 4 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 5 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 6 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 7 คะแนน

ฉบับที่ 5 แบบเดิมตัวเลข 8 มาตรา

		ข้อความทางบวก		ข้อความทางลบ	
เลือก	8	ให้คะแนน	8 คะแนน	ให้คะแนน	1 คะแนน
เลือก	7	ให้คะแนน	7 คะแนน	ให้คะแนน	2 คะแนน
เลือก	6	ให้คะแนน	6 คะแนน	ให้คะแนน	3 คะแนน
เลือก	5	ให้คะแนน	5 คะแนน	ให้คะแนน	4 คะแนน
เลือก	4	ให้คะแนน	4 คะแนน	ให้คะแนน	5 คะแนน
เลือก	3	ให้คะแนน	3 คะแนน	ให้คะแนน	6 คะแนน
เลือก	2	ให้คะแนน	2 คะแนน	ให้คะแนน	7 คะแนน
เลือก	1	ให้คะแนน	1 คะแนน	ให้คะแนน	8 คะแนน

ฉบับที่ 6 แบบเดิมตัวเลข 9 มาตรา

		ข้อความทางบวก		ข้อความทางลบ	
เลือก	9	ให้คะแนน	9 คะแนน	ให้คะแนน	1 คะแนน
เลือก	8	ให้คะแนน	8 คะแนน	ให้คะแนน	2 คะแนน
เลือก	7	ให้คะแนน	7 คะแนน	ให้คะแนน	3 คะแนน
เลือก	6	ให้คะแนน	6 คะแนน	ให้คะแนน	4 คะแนน
เลือก	5	ให้คะแนน	5 คะแนน	ให้คะแนน	5 คะแนน
เลือก	4	ให้คะแนน	4 คะแนน	ให้คะแนน	6 คะแนน
เลือก	3	ให้คะแนน	3 คะแนน	ให้คะแนน	7 คะแนน
เลือก	2	ให้คะแนน	2 คะแนน	ให้คะแนน	8 คะแนน
เลือก	1	ให้คะแนน	1 คะแนน	ให้คะแนน	9 คะแนน

ฉบับที่ 7 แบบบรรยาย 4 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก		ข้อความทางลบ	
เลือก	4	ให้คะแนน	4 คะแนน	ให้คะแนน	1 คะแนน
เลือก	3	ให้คะแนน	3 คะแนน	ให้คะแนน	2 คะแนน
เลือก	2	ให้คะแนน	2 คะแนน	ให้คะแนน	3 คะแนน
เลือก	1	ให้คะแนน	1 คะแนน	ให้คะแนน	4 คะแนน

ฉบับที่ 8 แบบบรรยาย 5 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก	ข้อความทางลบ
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 1 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 2 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 3 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 4 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 5 คะแนน

ฉบับที่ 9 แบบบรรยาย 6 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก	ข้อความทางลบ
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 1 คะแนน.
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 2 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 3 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 4 คะแนน.
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 5 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 6 คะแนน

ฉบับที่ 10 แบบบรรยาย 7 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก	ข้อความทางลบ
เลือก	7	ให้คะแนน 7 คะแนน	ให้คะแนน 1 คะแนน
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 2 คะแนน
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 3 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 4 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 5 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 6 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 7 คะแนน

ฉบับที่ 11 แบบบรรยาย 8 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก	ข้อความทางลบ
เลือก	8	ให้คะแนน 8 คะแนน	ให้คะแนน 1 คะแนน
เลือก	7	ให้คะแนน 7 คะแนน	ให้คะแนน 2 คะแนน
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 3 คะแนน
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 4 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 5 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 6 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 7 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 8 คะแนน

ฉบับที่ 12 แบบบรรยาย 9 มาตรา ให้คะแนนดังนี้

		ข้อความทางบวก	ข้อความทางลบ
เลือก	9	ให้คะแนน 9 คะแนน	ให้คะแนน 1 คะแนน
เลือก	8	ให้คะแนน 8 คะแนน	ให้คะแนน 2 คะแนน
เลือก	7	ให้คะแนน 7 คะแนน	ให้คะแนน 3 คะแนน
เลือก	6	ให้คะแนน 6 คะแนน	ให้คะแนน 4 คะแนน
เลือก	5	ให้คะแนน 5 คะแนน	ให้คะแนน 5 คะแนน
เลือก	4	ให้คะแนน 4 คะแนน	ให้คะแนน 6 คะแนน
เลือก	3	ให้คะแนน 3 คะแนน	ให้คะแนน 7 คะแนน
เลือก	2	ให้คะแนน 2 คะแนน	ให้คะแนน 8 คะแนน
เลือก	1	ให้คะแนน 1 คะแนน	ให้คะแนน 9 คะแนน

ฉบับที่ 13 แบบกราฟฟิก จำนวน 4 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก					
ไม่เห็นด้วยมากที่สุด	1	2	3	4	เห็นด้วยมากที่สุด
ข้อความทางลบ					
ไม่เห็นด้วยมากที่สุด	4	3	2	1	เห็นด้วยมากที่สุด

ฉบับที่ 14 แบบกราฟฟิก จำนวน 5 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก

ไม่เห็นด้วยมากที่สุด

1	2	3	4	5
---	---	---	---	---

 เห็นด้วยมากที่สุด

ข้อความทางลบ

ไม่เห็นด้วยมากที่สุด

5	4	3	2	1
---	---	---	---	---

 เห็นด้วยมากที่สุด

ฉบับที่ 15 แบบกราฟฟิก จำนวน 6 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก

ไม่เห็นด้วยมากที่สุด

1	2	3	4	5	6
---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ข้อความทางลบ

ไม่เห็นด้วยมากที่สุด

6	5	4	3	2	1
---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ฉบับที่ 16 แบบกราฟฟิก จำนวน 7 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก

ไม่เห็นด้วยมากที่สุด

1	2	3	4	5	6	7
---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ข้อความทางลบ

ไม่เห็นด้วยมากที่สุด

7	6	5	4	3	2	1
---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ฉบับที่ 17 แบบกราฟฟิก จำนวน 8 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก

ไม่เห็นด้วยมากที่สุด

1	2	3	4	5	6	7	8
---	---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ข้อความทางลบ

ไม่เห็นด้วยมากที่สุด

8	7	6	5	4	3	2	1
---	---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ฉบับที่ 18 แบบกราฟฟิก จำนวน 9 มาตรา กำหนดคะแนนโดยแบ่งเส้นกราฟดังนี้

ข้อความทางบวก

ไม่เห็นด้วยมากที่สุด

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

ข้อความทางลบ

ไม่เห็นด้วยมากที่สุด

9	8	7	6	5	4	3	2	1
---	---	---	---	---	---	---	---	---

 เห็นด้วยมากที่สุด

4. นำข้อมูลที่ได้มาวิเคราะห์ทางสถิติและการทดสอบสมมติฐานที่ตั้งไว้

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการศึกษานี้ ใช้การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์โดยใช้โปรแกรมสำเร็จรูป และเขียนคำสั่งเพิ่มเติมในกรณีที่ไม่สามารถวิเคราะห์ได้โดยตรง และผู้วิจัยได้ดำเนินการตามลำดับขั้น ดังนี้

1. ตรวจสอบความสมบูรณ์ของข้อมูลพร้อมรหัส (Code) บันทึกข้อมูลเก็บไว้ในเครื่องคอมพิวเตอร์
2. แบ่งมาตราประเมินค่าแต่ละฉบับจากจำนวน 27 ข้อ ออกเป็น 27 ส่วน (Parts) ด้วยความยาวที่เท่ากันได้ส่วนละ 1 ข้อ
3. หาค่าสถิติพื้นฐานของมาตราประเมินค่า ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวนร่วมระหว่างส่วนย่อยของมาตรวัด
4. คำนวณค่าความเชื่อมั่นด้วยสูตรหาค่าความเชื่อมั่นจากแบบจำลองคะแนนจริง สมมุติคือ สูตร α โดยใช้โปรแกรมสำเร็จรูป SPSS FOR WINDOW เวอร์ชัน 10 และคำนวณค่าความเชื่อมั่นจากแบบจำลองคะแนนจริงสัมพันธ์ คือ สูตร r_{F-R} โดยเขียนคำสั่งในโปรแกรม Excell
5. ทดสอบนัยสำคัญความแตกต่างของค่าความเชื่อมั่น โดยใช้สูตร UX_i (Wooddruft and Feldt. 1986) โดยใช้โปรแกรมสำเร็จ ALPHATEST และทดสอบความแตกต่างรายคู่ของค่าความเชื่อมั่น

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. หาค่าสถิติพื้นฐานของมาตราประเมีนค่า

1.1 ค่าเฉลี่ย

1.2 ค่าเบี่ยงเบนมาตรฐาน

2. วิเคราะห์ข้อสอบรายข้อ (Item Analysis) หาค่าอำนาจจำแนกของมาตราประเมินค่าโดยใช้สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Product-Moment Correlation) r_{xy} ซึ่งมีสูตรดังนี้ (บุญเชิด ภิญโญนนตพงษ์. 2545 : 165-166)

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

เมื่อ	r_{xy}	แทน	ค่าอำนาจจำแนก
	N	แทน	จำนวนคนในกลุ่ม
	X	แทน	คะแนนของข้อคำถาม
	Y	แทน	คะแนนผลรวมของข้ออื่นๆ ที่เหลือทุกข้อ

3. ค่าความเที่ยงตรงเชิงพินิจ โดยการพิจารณาค่าดัชนีความสอดคล้อง (IOC) (บุญเชิด ภิญโญนนตพงษ์. 2545 : 179)

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

เมื่อ	IOC	แทน	ดัชนีความสอดคล้องมีค่าอยู่ระหว่าง -1 ถึง +1
	$\sum R$	แทน	ผลรวมของการพิจารณาของผู้เชี่ยวชาญ
	N	แทน	จำนวนผู้เชี่ยวชาญ

4. สถิติที่ใช้คำนวณค่าความเชื่อมั่น โดยใช้สูตร r_{F-R} ของ Feldt- Raju (บุญเชิด ภิญโญนนตพงษ์.2545: 228)

$$r_{F-R} = \left[\frac{1}{1 - \sum \lambda_i^2} \right] \left[\frac{S_x^2 - \sum S_i^2}{S_x^2} \right]$$

เมื่อ	r_{F-R}	แทน	ความเชื่อมั่นของแบบทดสอบ
	λ_i	แทน	สัดส่วนระหว่างผลรวมของเมตริกในแถวที่ i กับความแปรปรวนของแบบทดสอบทั้งฉบับ
	S_i^2	แทน	ความแปรปรวนของส่วนย่อยของแบบทดสอบ
	S_x^2	แทน	ความแปรปรวนของแบบทดสอบทั้งฉบับ

5. สถิติที่ใช้คำนวณค่าความเชื่อมั่น โดยใช้สูตร สัมประสิทธิ์แอลฟา (α Coefficient) (บุญเชิด ภิญโญนนตพงษ์. 2545 : 220)

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

α	แทน	ค่าสัมประสิทธิ์ของความเชื่อมั่น
n	แทน	จำนวนข้อของเครื่องมือวัด
S_i^2	แทน	คะแนนความแปรปรวนเป็นรายข้อ
S_t^2	แทน	คะแนนความแปรปรวน ของเครื่องมือทั้งฉบับ

6. ทดสอบนัยสำคัญความแตกต่างของค่าความเชื่อมั่นโดยใช้ UX_1 (บุญเชิด ภิญโญนนตพงษ์. 2537 : 57 อ้างอิงจาก ; Wooddruff and Feldt. 1989 : 393 – 413. *Psychometika*.)

$$UX_1 = \sum_{i=1}^m \frac{[u_i - \bar{u}]^2}{S_u^2 - Cu}$$

$$\bar{u} = \sum_{i=1}^m \frac{u_i}{m}$$

$$u_1 = \frac{1}{[1 - r_1]^{1/3}}$$

$$S_u^2 = \frac{2}{9m[N_e - 1]} \sum_{i=1}^m u_i^2$$

$$C_u = \frac{4}{9m[m - 1][N_e - 1]} \sum_{i=2}^m \sum_{j=1}^{i-1} r_{ij}^2 u_i u_j$$

$$N_c = N \left[\frac{\bar{n}_k - 1}{\bar{n}_k + 1} \right]$$

$$\bar{n}_k = \frac{M}{\sum [1/n_i]}$$

เมื่อ	UX_i	แทน	สถิติทดสอบที่มีการแจกแจงแบบไค-กำลังสอง
	r	แทน	ค่าความเชื่อมั่นที่คำนวณได้
	r_{ij}	แทน	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร
	S_u^2	แทน	ความแปรปรวนของ u_i
	c_u	แทน	ความแปรปรวนร่วมของ u_i
	n	แทน	จำนวนค่าสังเกตหรือจำนวนตัวแปร
	m	แทน	จำนวนค่าความเชื่อมั่น

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้เกิดความเข้าใจตรงกันในการแปลความหมายผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้กำหนดสัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

α	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์แอลฟาของครอนบัค
r_{F-R}	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์ของเฟลด์ – ราชู
UX_1	แทน	สถิติทดสอบที่มีการแจกแจงแบบไค – กำลังสอง
n	แทน	จำนวนข้อในมาตรประเมินค่า
N	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
S	แทน	ส่วนเบี่ยงเบนมาตรฐาน

การเสนอผลการวิเคราะห์ข้อมูล

การศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับขั้นตอน ดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานของมาตรประเมินค่าทั้ง 18 ฉบับ

ตอนที่ 2 ค่าความเชื่อมั่นของมาตรประเมินค่าทั้ง 18 ฉบับ

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบต่างกัน คือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จำแนกตามจำนวนมาตรา

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มาตรา จำแนกตามรูปแบบมาตรประเมินค่า

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ค่าสถิติพื้นฐานของมาตรประเมินค่าทั้ง 18 ฉบับ

การวิเคราะห์ข้อมูลตอนนี้ ผู้วิจัยได้นำคะแนนมาตรประเมินค่าทั้ง 3 รูปแบบ คือ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก โดยในแต่ละรูปแบบมีจำนวน 4, 5, 6, 7, 8 และ 9 มาตรา รวมทั้งหมด 18 ฉบับ มาคำนวณหาค่าเฉลี่ยและความเบี่ยงเบนมาตรฐาน ดังแสดงในตาราง 2

ตาราง 2 ค่าเฉลี่ย ($\bar{X}$) และความเบี่ยงเบนมาตรฐาน (S.D.) ของคะแนนทั้งฉบับของ
 มาตรฐานค่า 3 รูปแบบ จำแนกตามจำนวนมาตรา

จำนวนมาตรา	รูปแบบ					
	เติมตัวเลข		บรรยาย		กราฟฟิก	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
4 มาตรา	74.8222	6.9728	73.6778	8.9056	73.7222	10.3598,
5 มาตรา	91.8667	8.8878	84.0222	14.1834	89.6444	11.8230,
6 มาตรา	106.9444	12.3492	105.5667	14.1295	103.5000	18.9738,
7 มาตรา	123.9222	14.2271	110.8556	23.2841	122.4667	17.5967,
8 มาตรา	144.2222	21.3688	141.6111	26.5883	135.2556	26.1339
9 มาตรา	160.5222	20.6899	139.1778	32.0097	153.3778	21.6984

ผลการวิเคราะห์ตามตาราง 2 พบว่า มาตรฐานค่าแบบเติมตัวเลข 4, 5, 6, 7, 8 และ 9 มาตรา มีค่าเฉลี่ยของคะแนนทั้งฉบับเท่ากับ 74.8222, 91.8667, 106.9444, 123.9222, 144.2222 และ 160.5222 ตามลำดับ และส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งฉบับมีค่าเท่ากับ 6.9728, 8.8878, 12.3492, 14.2271, 21.3688 และ 20.6899 ตามลำดับ

มาตรฐานค่าแบบบรรยาย 4, 5, 6, 7, 8 และ 9 มาตรา มีค่าเฉลี่ยของคะแนนทั้งฉบับเท่ากับ 73.6778, 84.0222, 105.5667, 110.8556, 141.6111 และ 139.1778 ตามลำดับ และส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งฉบับมีค่าเท่ากับ 8.9056, 14.1834, 14.1295, 23.2841, 26.5883 และ 32.0097 ตามลำดับ

มาตรฐานค่าแบบกราฟฟิก 4, 5, 6, 7, 8 และ 9 มาตรา มีค่าเฉลี่ยของคะแนนทั้งฉบับเท่ากับ 73.7222, 89.6444, 103.5000, 122.4667, 135.2556 และ 153.3778 ตามลำดับ และส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งฉบับมีค่าเท่ากับ 10.3598, 11.8230, 18.9738, 17.5967, 26.1339 และ 21.6984 ตามลำดับ

ตอนที่ 2 ค่าความเชื่อมั่นของมาตรประเมินค่า

ในการวิเคราะห์ข้อมูลตอนนี้ ผู้วิจัยได้นำคะแนนของมาตรประเมินค่าทั้งสามรูปแบบ ที่มีจำนวนมาตรต่างกัน ทั้งหมด 18 ฉบับ มาหาค่าความแปรปรวนในแต่ละส่วนย่อย 27 ส่วน และความแปรปรวนทั้งฉบับ แล้วคำนวณค่าความเชื่อมั่นโดยใช้สูตร α และ r_{F-R} ปรากฏ ดังแสดงในตาราง 3

ตาราง 3 ความเชื่อมั่นของมาตรประเมินค่าทั้ง 18 ฉบับ คำนวณโดยใช้สูตร α และ r_{F-R}

จำนวนมาตรา	รูปแบบมาตรประเมินค่า					
	เติมตัวเลข		บรรยาย		กราฟฟิก	
	α	r_{FR}	α	r_{FR}	α	r_{FR}
4 มาตรา	.5845	.5884	.7946	.7997	.7972	.7985
5 มาตรา	.6402	.6470	.8595	.8613	.7582	.7645
6 มาตรา	.6646	.6683	.8038	.8063	.8362	.8383
7 มาตรา	.6695	.6801	.8878	.8896	.7806	.7847
8 มาตรา	.7718	.7755	.8609	.8826	.8447	.8463
9 มาตรา	.7137	.7215	.8891	.8564	.7411	.7459

ผลการวิเคราะห์ตามตาราง 3 ได้ดังนี้

1. เมื่อใช้สูตร α คำนวณ พบว่า

มาตรประเมินค่าแบบเติมตัวเลขและแบบกราฟฟิก จำนวน 8 มาตรา มีค่าความเชื่อมั่นสูงสุด (.7718 และ .8463 ตามลำดับ) ส่วนแบบบรรยาย จำนวน 9 มาตรา มีค่าความเชื่อมั่นสูงสุด (.8891)

มาตรประเมินค่าจำนวน 4 มาตรา แบบกราฟฟิกมีค่าความเชื่อมั่นสูงสุด (.7972) มาตรประเมินค่าจำนวน 5 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8595) มาตรประเมินค่าจำนวน 6 มาตรา แบบกราฟฟิกมีค่าความเชื่อมั่นสูงสุด (.8362) มาตรประเมินค่าจำนวน 7 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8878) มาตรประเมินค่าจำนวน 8 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8609) มาตรประเมินค่าจำนวน 9 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8891)

2. เมื่อใช้สูตร $r_{F.R}$ คำนวณ พบว่า

มาตรประเมินค่าแบบเติมตัวเลขและแบบกราฟฟิก จำนวน 8 มาตรา มีค่าความเชื่อมั่นสูงสุด (.7755 และ .8463) ส่วนแบบบรรยาย จำนวน 7 มาตรา มีค่าความเชื่อมั่นสูงสุด (.8896)

มาตรประเมินค่าจำนวน 4 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.7997) มาตรประเมินค่าจำนวน 5 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8613) มาตรประเมินค่าจำนวน 6 มาตรา แบบกราฟฟิกมีค่าความเชื่อมั่นสูงสุด (.8383) มาตรประเมินค่าจำนวน 7 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8896) มาตรประเมินค่า จำนวน 8 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8826) มาตรประเมินค่าจำนวน 9 มาตรา แบบบรรยายมีค่าความเชื่อมั่นสูงสุด (.8564)

ตอนที่ 3 การเปรียบเทียบความเชื่อมั่นของมาตรประเมินค่าที่มีรูปแบบต่างกัน จำแนกตามจำนวนมาตรา

3.1 การวิเคราะห์ตอนนี้ผู้วิจัยได้นำค่าความเชื่อมั่นของมาตรประเมินค่า ที่มีรูปแบบต่างกัน 3 ฉบับ คือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก มาทดสอบความแตกต่างกัน โดยใช้สูตร UX_1 จำแนกตามจำนวนมาตรา ปรากฏผลดังแสดงในตาราง 4

ตาราง 4 เปรียบเทียบค่าความเชื่อมั่นของมาตรวัดที่มีรูปแบบต่างกัน 3 รูปแบบ จำแนกตามจำนวนมาตรา

จำนวนมาตรา		ค่าความเชื่อมั่นของมาตรประเมินค่า			UX ₁
		เติมตัวเลข	บรรยาย	แบบกราฟฟิก	
4 มาตรา	α	.5845	.7946	.7972	12.6174**
	r _{FR}	.5884	.7997	.7985	12.8277**
5 มาตรา	α	.6402	.8595	.7582	18.4114**
	r _{FR}	.6470	.8613	.7645	18.0810**
6 มาตรา	α	.6646	.8038	.8362	10.7191**
	r _{FR}	.6683	.8063	.8383	10.7736**
7 มาตรา	α	.6695	.8878	.7806	24.8150**
	r _{FR}	.6801	.8896	.7847	24.1793**
8 มาตรา	α	.7718	.8609	.8447	5.3192
	r _{FR}	.7755	.8826	.8463	5.2182
9 มาตรา	α	.7137	.8891	.7411	23.8355**
	r _{FR}	.7215	.8564	.7459	11.1605**

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 4 ได้ดังนี้

1. เมื่อใช้สูตร α คำนวณ พบว่า ค่าความเชื่อมั่นของมาตรประเมินค่าแบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในเกือบทุกจำนวนมาตรา ยกเว้น เมื่อเป็นจำนวน 8 มาตรา เท่านั้นที่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. เมื่อใช้สูตร r_{F-R} คำนวณ พบว่า ค่าความเชื่อมั่นของมาตรประเมินค่าแบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในเกือบทุกจำนวนมาตรา ยกเว้น เมื่อเป็นจำนวน 8 มาตรา เท่านั้นที่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.2 การวิเคราะห์ตอนนี้เป็นการเปรียบเทียบรายคู่ ผู้วิจัยได้นำค่าความเชื่อมั่นรายคู่ระหว่างมาตรฐานประเมินค่าที่มีรูปแบบต่างกัน เมื่อมีจำนวนมาตราเท่ากันมาเปรียบเทียบกันและทดสอบความแตกต่างกันโดยใช้สูตร UX_1 ปรากฏผลดังแสดงในตาราง 5

ตาราง 5 เปรียบเทียบความแตกต่างค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าที่มีรูปแบบต่างกัน จำแนกตามจำนวนมาตรา

สูตร	จำนวน มาตรา	UX_1 ของการเปรียบเทียบรายคู่แบบ		
		เต็มตัวเลข-บรรยาย	เต็มตัวเลข-กราฟฟิก	บรรยาย-กราฟฟิก
α	4 มาตรา	10.0159**	10.3728**	.0033
	5 มาตรา	17.5347**	3.2370	6.0024**
	6 มาตรา	10.0159**	10.3728**	.0033
	7 มาตรา	22.8557**	3.4384	9.0941**
	9 มาตรา	22.8557**	3.4384	9.0941**
r_{F-R}	4 มาตรา	10.4570**	10.2883**	.0007
	5 มาตรา	17.3137**	3.3564	5.7113*
	6 มาตรา	10.4570**	10.2883**	.0007
	7 มาตรา	22.2004**	3.2134	9.0227**
	9 มาตรา	22.2004**	3.2134	9.0227**

* มีนัยสำคัญทางสถิติที่ระดับ .05

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 5 ได้ดังนี้

1. เมื่อใช้สูตร α คำนวณ พบว่า

มาตรฐานประเมินค่าแบบเติมตัวเลขกับแบบบรรยายมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกจำนวนมาตรา

มาตรฐานประเมินค่าแบบเติมตัวเลขกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 4 และ 5 มาตรา

มาตรฐานประเมินค่าแบบบรรยายกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 5, 7 และ 9 มาตรา

2. เมื่อใช้สูตร r_{F-R} คำนวณ พบว่า

มาตรฐานประเมินค่าแบบเติมตัวเลขกับแบบบรรยายมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกจำนวนมาตรา

มาตรฐานประเมินค่าแบบเติมตัวเลขกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 4 และ 5 มาตรา

มาตรฐานประเมินค่าแบบบรรยายกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และที่ระดับ .01 เฉพาะที่จำนวนมาตรา 5, 7 และ 9 มาตรา ตามลำดับ

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวน
มาตราต่างกัน จำแนกตามรูปแบบมาตรประเมินค่า

4.1 การวิเคราะห์ตอนนี้ได้นำค่าความเชื่อมั่นของมาตรการประเมินค่าที่มีจำนวน
มาตราต่างกัน ทั้ง 6 ฉบับ คือ 4, 5, 6, 7, 8 และ 9 มาตรา ความแตกต่างโดยใช้สูตร UX_1
จำแนกตามรูปแบบมาตรประเมินค่า ปรากฏผลดังแสดงในตาราง 6

ตาราง 6 เปรียบเทียบค่าความเชื่อมั่นของมาตรประเมินค่าที่มีจำนวนมาตราต่างกัน จำแนก
ตามตามรูปแบบมาตรประเมินค่า

รูปแบบ		มาตรประเมินค่าจำนวน						UX_1
		4	5	6	7	8	9	
		มาตรา	มาตรา	มาตรา	มาตรา	มาตรา	มาตรา	
เติมตัวเลข	α	.5845	.6402	.6646	.6695	.7718	.7137	9.0101
	r_{F-R}	.5884	.6470	.6683	.6801	.7755	.7215	9.1911
บรรยาย	α	.7946	.8595	.8038	.8878	.8609	.8891	14.1010*
	r_{F-R}	.7997	.8613	.8063	.8896	.8826	.8564	10.5166
กราฟฟิก	α	.7972	.7582	.8362	.7806	.8447	.7411	8.9087
	r_{F-R}	.7985	.7645	.8383	.7847	.8463	.7459	8.5197

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 6 ได้ดังนี้

- เมื่อใช้สูตร α คำนวณ พบว่า มาตรประเมินค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เฉพาะมาตรประเมินค่าแบบบรรยายเท่านั้น
- เมื่อใช้สูตร r_{F-R} คำนวณ พบว่า มาตรประเมินค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีความนัยสำคัญทางสถิติ ทุกรูปแบบ

4.2 การวิเคราะห์ตอนนี้เป็นการเปรียบเทียบรายคู่ ผู้วิจัยได้นำค่าความเชื่อมั่นรายคู่ระหว่างมาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกัน เมื่อมีรูปแบบบรรยายมาเปรียบเทียบกันและทดสอบความแตกต่างกันโดยใช้สูตร UX_1 ปรากฏผลดังแสดงในตาราง 7

ตาราง 7 เปรียบเทียบความแตกต่างค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าแบบบรรยายที่มีจำนวนมาตราต่างกัน

ค่า UX_1 ของการเปรียบเทียบรายค่า α รายคู่					
	5 มาตรา	6 มาตรา	7 มาตรา	8 มาตรา	9 มาตรา
4 มาตรา	2.9572	.0433	7.4220**	3.141	7.7059**
5 มาตรา		2.2900	1.0419	.0021	1.1524
6 มาตรา			.3552**	2.4287	6.6190**
7 มาตรา				.9514	.0028
8 มาตรา					1.0571

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 7 ได้ดังนี้

เมื่อใช้สูตร α คำนวณ พบว่า ค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าแบบบรรยายแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เพียง 4 คู่ ได้แก่ จำนวนมาตรา 4 มาตรากับ 7 มาตรา 4 มาตรากับ 9 มาตรา 6 มาตรากับ 7 มาตรา และ 6 มาตรากับ 9 มาตรา ส่วนคู่อื่นนอกจากนี้แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

จุดมุ่งหมายของการศึกษาค้นคว้า

จุดมุ่งหมายสำคัญของการศึกษาค้นคว้าครั้งนี้ เพื่อศึกษาว่ามาตรฐานประเมินค่าที่มีรูปแบบมาตรวัดต่างกัน 3 รูปแบบ ได้แก่ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก รูปแบบใดมีค่าความเชื่อมั่นมากที่สุด และจะมีค่าแตกต่างกันหรือไม่ เมื่อจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มาตรา

โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบค่าความเชื่อมั่นของมาตรฐานประเมินค่าที่มีรูปแบบต่างกัน คือ แบบเติมตัวเลข แบบบรรยาย และแบบกราฟฟิก จำแนกตามจำนวนมาตรา
2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของมาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกัน คือ 4, 5, 6, 7, 8 และ 9 มาตรา จำแนกตามรูปแบบมาตรฐานประเมินค่า

สมมติฐานของการวิจัย

1. มาตรฐานที่มีรูปแบบมาตรวัดแตกต่างกัน คือ แบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก มีค่าความเชื่อมั่นแตกต่างกัน
2. มาตรฐานประเมินค่าแต่ละรูปแบบเดียวกันที่มีจำนวนมาตราต่างกัน 6 มาตรา คือ 4, 5, 6, 7, 8 และ 9 มาตรา มีค่าความเชื่อมั่นแตกต่างกัน

ประชากรและกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ของโรงเรียนมัธยมศึกษาในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2545 ของโรงเรียนมัธยมศึกษาในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 6 โรงเรียน จำนวน 540 คน ซึ่งได้มาโดยวิธีสุ่มหลายขั้นตอน (Multistage Random Sampling)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นมาตรฐานประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบ ซึ่งผู้วิจัยนำมาตรฐานประเมินค่าลักษณะนิสัยการทำให้ตนเองเสียเปรียบของ จอห์น และโรดวอลท์ (Jones And Rhodewalt. 1982) มาแปลและเรียบเรียงเป็นภาษาไทย จำนวน 27 ข้อ ซึ่งแบ่งเป็น 27 ส่วน ส่วนละ 1 ข้อ โดยผู้วิจัยได้ดัดแปลงรูปแบบการตอบของมาตรวัดเป็น 3 รูปแบบ คือ มาตรฐานประเมินค่าแบบเต็มตัวเลข แบบกราฟฟิก และแบบบรรยาย โดยทั้ง 3 รูปแบบมีจำนวนมาตรา 4 มาตรา 5 มาตรา 6 มาตรา 7 มาตรา 8 มาตรา และ 9 มาตรา

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการศึกษาค้นคว้าครั้งนี้ ใช้การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์โดยใช้โปรแกรมสำเร็จรูป และเขียนคำสั่งเพิ่มเติมในกรณีที่ไม่สามารถวิเคราะห์ได้โดยตรง และผู้วิจัยได้ดำเนินการตามลำดับขั้น ดังนี้

1. ตรวจสอบความสมบูรณ์ของข้อมูลพร้อมลงรหัส (Code) บันทึกข้อมูลเก็บไว้ในเครื่องคอมพิวเตอร์
2. แบ่งมาตรฐานประเมินค่าแต่ละฉบับจากจำนวน 27 ข้อ ออกเป็น 27 ส่วน (Parts) ด้วยความยาวที่เท่ากันได้ส่วนละ 1 ข้อ
3. หาค่าสถิติพื้นฐานของมาตรฐานประเมินค่า ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน
4. คำนวณค่าความเชื่อมั่นด้วยสูตรหาค่าความเชื่อมั่นจากแบบจำลองคะแนนจริง
สมมุติคือ สูตร α โดยใช้โปรแกรมสำเร็จรูป SPSS FOR WINDOW เวอร์ชัน 10 และคำนวณค่าความเชื่อมั่นด้วยสูตรหาค่าความเชื่อมั่นจากแบบจำลองคะแนนจริงสัมพันธ์ คือ สูตร r_{F-R} โดยเขียนคำสั่งในโปรแกรม Excell
5. ทดสอบนัยสำคัญความแตกต่างของค่าความเชื่อมั่น โดยใช้สูตร UX_1 (Wooddruff

and Feldt. 1986) โดยใช้โปรแกรม ALPHATEST และทดสอบความแตกต่างรายคู่ของค่าความเชื่อมั่น

สรุปผลการวิเคราะห์ข้อมูล

ผลการศึกษาค้นคว้า สรุปได้ดังนี้

1. ผลการศึกษาค่าความเชื่อมั่นของมาตรประเมินค่าทั้ง 18 ฉบับ เมื่อใช้สูตร α คำนวณพบว่ามาตรประเมินค่าแบบเติมตัวเลขและแบบกราฟฟิก จำนวน 8 มาตรามีค่าความเชื่อมั่นสูงสุด ส่วนแบบบรรยาย จำนวน 9 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 4 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 5 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 6 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 7 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 8 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 9 มาตรามีค่าความเชื่อมั่นสูงสุด

เมื่อใช้สูตร r_{FR} คำนวณพบว่า มาตรประเมินค่าแบบเติมตัวเลขและแบบกราฟฟิก จำนวน 8 มาตรามีค่าความเชื่อมั่นสูงสุด ส่วนแบบบรรยาย จำนวน 7 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 4 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 5 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 6 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 7 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 8 มาตรามีค่าความเชื่อมั่นสูงสุด มาตรประเมินค่าจำนวน 9 มาตรามีค่าความเชื่อมั่นสูงสุด

2. ผลการเปรียบเทียบค่าความเชื่อมั่นมาตรประเมินค่าที่มีรูปแบบแตกต่างกัน คือแบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก จำแนกตามจำนวนมาตรา เมื่อใช้สูตร α และ r_{FR} คำนวณพบว่า ค่าความเชื่อมั่นมาตรประเมินค่าแบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในเกือบทุกจำนวนมาตรา ยกเว้น เมื่อเป็นจำนวน 8 มาตรานั้นที่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จากนั้นทดสอบค่าความเชื่อมั่นรายคู่ของมาตรประเมินค่าที่คำนวณโดยใช้สูตร α พบว่า มาตรประเมินค่าที่มีรูปแบบเติมตัวเลขกับแบบบรรยายมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกจำนวนมาตรา มาตรประเมินค่าที่มีรูปแบบเติมตัวเลขกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 4 และ 5 มาตรานั้น และมาตรประเมินค่าแบบบรรยายกับแบบกราฟฟิกมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 5, 7 และ 9 มาตรานั้น

และเมื่อทดสอบค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าที่คำนวณโดยใช้สูตร r_{F-R} ปรากฏว่าให้ผลเช่นเดียวกัน ยกเว้นคู่ของมาตรฐานประเมินค่าแบบบรรยายกับแบบกราฟฟิคมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เฉพาะที่จำนวนมาตรา 5 มาตรา เท่านั้น และแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เฉพาะที่จำนวนมาตรา 7 และ 9 มาตรา เท่านั้น

3. ผลการเปรียบเทียบค่าความเชื่อมั่นของมาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกัน 6 มาตรา คือ จำนวนมาตรา 4, 5, 6, 7, 8 และ 9 มาตรา จำแนกตามรูปแบบมาตรฐานประเมินค่า เมื่อใช้สูตร α คำนวณ พบว่ามาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เฉพาะมาตรฐานประเมินค่าแบบบรรยายเท่านั้น เมื่อใช้สูตร r_{F-R} คำนวณ พบว่า มาตรฐานประเมินค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทุกรูปแบบ จากนั้นทดสอบค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าแบบบรรยายที่คำนวณโดยใช้สูตร α พบว่า ค่าความเชื่อมั่นรายคู่ของมาตรฐานประเมินค่าแบบบรรยายแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เพียง 4 คู่ ได้แก่ จำนวนมาตรา 4 มาตรากับ 7 มาตรา 4 มาตรากับ 9 มาตรา 6 มาตรากับ 7 มาตรา และ 6 มาตรากับ 9 มาตรา ส่วนคู่อื่นนอกจากนี้แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

ผลการศึกษาค้นคว้า อภิปรายได้ดังนี้

1. ค่าความเชื่อมั่นมาตรฐานประเมินค่า ที่มีรูปแบบแตกต่างกัน คือ แบบเติมตัวเลขแบบบรรยาย แบบกราฟฟิค จำแนกตามจำนวนมาตรา เมื่อใช้สูตร α และ r_{F-R} คำนวณ พบว่าค่าความเชื่อมั่นมาตรฐานประเมินค่าแบบเติมตัวเลข แบบบรรยาย แบบกราฟฟิค แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในเกือบทุกจำนวนมาตรา ยกเว้น เมื่อเป็นจำนวน 8 มาตรา เท่านั้นที่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งโดยภาพรวมแล้วเป็นไปตามสมมติฐานที่ตั้งไว้ ซึ่งสอดคล้องกับงานวิจัยของ พนิดา พุ่งเกียรติินาสุข. (2540 : 72) ที่ได้ศึกษาการประมาณค่าความเชื่อมั่นของมาตรฐานที่มีรูปแบบการตอบแตกต่างกัน 3 รูปแบบคือ แบบตัวเลขแบบบรรยาย แบบกราฟฟิค พบว่า มาตรฐานที่มีรูปแบบการตอบแตกต่างกัน มีค่าความเชื่อมั่นแตกต่างกัน เมื่อทดสอบค่าความเชื่อมั่นรายคู่ พบว่า มาตรฐานประเมินค่าแบบบรรยายมีค่าความเชื่อมั่นสูงกว่ามาตรฐานประเมินค่าแบบเติมตัวเลข ทุกจำนวนมาตรา และสูงกว่าแบบกราฟฟิค เฉพาะที่จำนวนมาตรา 5, 7 และ 9 มาตรา ตามลำดับ ซึ่งสอดคล้องกับงานวิจัยของ ปีเตอร์ และ แมคคอมมิคซ์ (Peter and McCormic.1996 : abstract) ที่ได้ทำการเปรียบเทียบความเชื่อมั่นแบบสอดคล้องของคำตอบจากผู้ประเมินของมาตรฐานประเมินค่าที่มีรูปแบบคำตอบต่างกัน 2 แบบ

คือ แบบตัวเลขและแบบบรรยาย พบว่าแบบบรรยายมีค่าความเชื่อมั่นสูงกว่าแบบตัวเลข อย่างมีนัยสำคัญ ซึ่งอาจเป็นเพราะว่า มาตรฐานค่าแบบบรรยาย เป็นมาตรวัดที่ผู้ตอบสามารถเข้าใจและเลือกได้ง่าย เพราะผู้ตอบไม่ต้องไปดูความหมายของสเกลอีก และยังเป็นการแก้ปัญหาสำคัญของผู้ที่ไม่เข้าใจสิ่งที่เป็นนามธรรม นอกจากนี้ มาตรฐานค่าแบบบรรยายยังเป็นมาตรวัดที่มีลักษณะเชิงพฤติกรรมของแต่ละลักษณะ ทำให้ผู้ตอบแต่ละคนมีความเข้าใจตรงกันว่าแต่ละระดับมีความมากน้อยเพียงใด สำหรับนักเรียนระดับมัธยมศึกษา นักเรียนชอบในสิ่งที่เข้าใจง่ายและสะดวก ดังนั้น มาตรฐานค่าแบบบรรยายจึงเหมาะกับนักเรียนมากที่สุด ในขณะที่ แบบเติมตัวเลข ผู้ตอบต้องนำตัวเลขไปเขียนตอบและไม่สะดวกสำหรับผู้ตอบที่ต้องกลับไปดูความหมายของตัวเลขก่อนตอบจึงอาจทำให้ผู้ตอบเกิดความสับสนได้ จึงเป็นแบบที่ทำให้เกิดความคลาดเคลื่อนได้มาก ส่วนแบบกราฟฟิกที่มีคำอธิบายหัวท้าย มีข้อเสียคือ ผู้ตอบมักจะเลือกเฉพาะปลาย 2 ข้างที่มีการอธิบาย (Thorndike. 1977 : 153) โดยเฉพาะผู้ที่ไม่เข้าใจในสิ่งที่เป็นนามธรรม (Wells and Smith. 1960 : abstract) ดังนั้น มาตรฐานค่าแบบกราฟฟิก จึงไม่ใช่มาตรฐานค่าที่เหมาะสมกับนักเรียนเป็นผู้ประเมิน

2. ค่าความเชื่อมั่นของมาตรฐานค่าที่มีจำนวนมาตราต่างกัน 6 มาตร คือ จำนวนมาตรา 4, 5, 6, 7, 8 และ 9 มาตร จำแนกตามรูปแบบมาตรฐานค่า เมื่อใช้สูตร α คำนวณ พบว่ามาตรฐานค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เฉพาะมาตรฐานค่าแบบบรรยายเท่านั้น เมื่อใช้สูตร r_{F-R} คำนวณ พบว่า มาตรฐานค่าที่มีจำนวนมาตราต่างกันมีค่าความเชื่อมั่นแตกต่างกัน อย่างไม่มีนัยสำคัญทางสถิติ ทุกรูปแบบ แสดงว่า มาตรฐานค่าที่มีจำนวนมาตราต่างกัน จำแนกเกือบทุกรูปแบบ มีแนวโน้มไม่แตกต่างกัน ทั้งนี้อาจเป็นเพราะว่าการวิจัยครั้งนี้ผู้วิจัยได้เก็บข้อมูล โดยให้นักเรียนแต่ละโรงเรียนตอบมาตรฐานค่าที่มีรูปแบบเดียวกันซ้ำกัน ตามมาตรฐานค่าที่มีจำนวนมาตรา 4, 6 และ 8 มาตร หรือที่มีจำนวน 5, 7 และ 9 มาตร ตามรูปแบบการเก็บรวบรวมข้อมูลดังกล่าวนี้อาจส่งผลกระทบต่อให้นักเรียนได้รับอิทธิพลการตอบ รูปแบบมาตรฐานค่าในฉบับแรก ทำให้ผลการตอบในฉบับต่อมามีคำตอบที่คล้ายคลึงกันมากกว่า ที่ควรจะเป็น นอกจากนั้นอาจมีผลด้านความเหนื่อยล้า เบื่อ หรือ อาจไม่เต็มใจตอบ ทำให้คำตอบที่ได้มีค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งฉบับของมาตรฐานค่าที่มี 4, 6 และ 8 มาตร หรือที่มี 5, 7 และ 9 มาตร ในแต่ละรูปแบบ ต่างกันน้อยมาก ส่งผลให้ค่าความเชื่อมั่นมีค่าพอ ๆ กัน ทำให้ไม่เป็นไปตามสมมติฐานที่ตั้งไว้ ซึ่งผลการวิจัยครั้งนี้มีความสอดคล้องกับผลการวิจัยของ ธารทิพย์ ประเสริฐสม (2522 : บทคัดย่อ) ที่ได้ศึกษาค่าความเชื่อมั่นของมาตรฐานค่าที่มีจำนวนช่วงมาตรตั้งแต่ 2 – 7 ช่วง โดยมีเครื่องมือที่ใช้วัดคือ แบบวัดทัศนคติต่อวิชาชีพครู และแบบวัดทัศนคติต่อการสัมภาษณ์เพื่อคัดเลือกการศึกษาคู พบว่า ความเชื่อมั่นของมาตรวัดทัศนคติต่อวิชาชีพครูมีจำนวนช่วงมาตร 3 – 7 ช่วง ไม่มีความแตกต่างกัน ส่วนค่าความเชื่อมั่นของแบบวัดทัศนคติต่อการสัมภาษณ์เพื่อคัดเลือกนักศึกษาครูที่

มีจำนวนช่วงมาตร 2 – 7 ช่วง ไม่มีความแตกต่างกัน ซึ่งสอดคล้องกับผลการวิจัยของ เบนดิง (Bending. 1953 : abstract) ซึ่งได้ศึกษาความเชื่อมั่นของมาตรประเมินค่าที่มีรายการคำตอบ เป็น 3 – 9 รายการ พบว่า ไม่มีความแตกต่างกัน เช่นเดียวกับการศึกษาของ วิยะดา หอธรรมอนันต์ (2526 : บทคัดย่อ) ที่ได้เปรียบเทียบความเชื่อมั่นแบบสอบซ้ำของมาตรประเมินค่าที่มีจำนวนช่วง 3 – 7 ช่วง โดยเว้นช่วงเวลาระหว่างการทดสอบครั้งที่ 1 และครั้งที่ 2 เป็น 1, 2, 3 และ 4 สัปดาห์ ผลการวิจัยพบว่า ค่าความเชื่อมั่นภายในของมาตรประเมินค่าในการทดสอบครั้งที่ 1 มีค่าความเชื่อมั่นที่ได้จากมาตรวัดที่มีจำนวนช่วง 3 – 7 ช่วง ไม่มีความแตกต่างกัน ในการทดสอบครั้งที่ 2 ค่าความเชื่อมั่นแบบคงที่ภายในของมาตรวัด ที่มีจำนวนช่วง 3 – 7 ช่วง ไม่มีความแตกต่างกัน

ข้อเสนอแนะในการวิจัย

1. ข้อเสนอแนะในการนำไปใช้

1.1 เนื่องจากมาตรประเมินค่าแบบบรรยายมีค่าความเชื่อมั่นสูงสุดในเกือบทุกจำนวนมาตราเมื่อเทียบกับมาตรประเมินค่ารูปแบบอื่น ดังนั้นในการศึกษาค้นคว้าผู้วิจัยควรเลือกใช้มาตรประเมินค่าแบบบรรยาย

1.3 เนื่องจากจำนวนมาตราส่งผลต่อค่าความเชื่อมั่นน้อยมาก ดังนั้นหากต้องตัดสินใจเลือกมาตรประเมินค่าว่าควรมีกี่จำนวนมาตราจึงขึ้นอยู่กับความต้องการของผู้ทำการวิจัยว่าต้องการให้แต่ละรูปแบบของมาตรประเมินค่า นั้น ละเอียดมากน้อยเพียงไร ถ้าต้องการคำตอบที่หยาบ ก็อาจใช้จำนวนมาตราน้อย ๆ แต่ถ้าต้องการคำตอบที่ละเอียดก็อาจเพิ่มจำนวนมาตรา เนื่องจาก จากผลการวิจัยพบว่าจำนวนมาตราแทบมีผลต่อการตอบของผู้ตอบไม่ต่างกัน

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากการวิจัยครั้งนี้ได้นำสูตร α และ r_{F-R} มาคำนวณค่าความเชื่อมั่นในลักษณะค่าบ่งชี้ โดยมีได้ทำการศึกษเปรียบเทียบ ดังนั้นควรมีการศึกษาเปรียบเทียบสูตรทั้งสองโดยแบ่งส่วนย่อยในลักษณะต่างกัน

2.2 ผลการวิจัยครั้งนี้ยังมิได้ข้อสรุปที่ชัดเจนว่าจำนวนมาตราต่างกันมีผลต่อค่าความเชื่อมั่น เนื่องจากการออกแบบการเก็บรวบรวมข้อมูลเป็นแบบเก็บซ้ำจำนวนมาตรา เพื่อให้ได้ข้อสรุปที่ชัดเจนควรมีการศึกษาวิจัยซ้ำโดยออกแบบการเก็บรวบรวมข้อมูลที่ใช้กลุ่มตัวอย่างอิสระเปรียบเทียบกับกลุ่มตัวอย่างที่สอบถามซ้ำ

2.3 การวิจัยครั้งนี้ได้นำมาวัดประเมินค่าแบบเดิมตัวเลขมาศึกษาเปรียบเทียบกับแบบบรรยายและแบบกราฟฟิค ซึ่งพบว่ามีค่าความเชื่อมั่นต่ำกว่าแบบอื่น ๆ โดยมีได้พิจารณาถึงต้นทุนการใช้ เช่นค่าใช้จ่ายและเวลาการตอบ ดังนั้นจึงควรนำประเด็นดังกล่าวนี้มาศึกษาเปรียบเทียบให้ละเอียดลุ่มลึกต่อไป

บรรณานุกรม

บรรณานุกรม

- กาญจนา มณีแสง. (2523). การประเมินผลทางการศึกษา. กรุงเทพฯ : ภาควิชาพื้นฐานการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ บางแสน.
- เกศเกล้า หมุ่มมิ่ง. (2539). การประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่า. ปรินูญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- ฉวีวรรณ ครองยุทธ. (2539). การประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่าที่มีขนาดกลุ่มตัวอย่างต่างกัน. ปรินูญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- ชูศรี วงศ์รัตน์. (2528). แบบแผนการทดลองและสถิติ. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- (2541). เทคนิคการใช้สถิติเพื่อการวิจัย. พิมพ์ครั้งที่ 7. กรุงเทพฯ : เทพเนรมิตรการพิมพ์.
- เชิดศักดิ์ โฉวาสินธุ์. (2520). การวัดทัศนคติและบุคลิกภาพ. กรุงเทพฯ : สำนักทดสอบทางการศึกษา และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ .
- (2525). การวัดผลทางการศึกษา. กรุงเทพฯ : สำนักทดสอบทางการศึกษา และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- ธารทิพย์ ประเสริฐสม. (2522). การเปรียบเทียบค่าความเที่ยงตรงของมาตรวัดทัศนคติแบบลิเคิร์ตที่มีจำนวนรายการคำตอบต่างกัน. วิทยานิพนธ์ ค.ม. (การวิจัยทางการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- นำบุญ เจนสรรพกิจกุล. (2534). การเปรียบเทียบผลการตอบมาตรฐานประเมินค่าที่มีรูปแบบคำตอบและลักษณะคำถามต่างกัน. วิทยานิพนธ์ ค.ม. (การวิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- บวร คมสิริพิทักษ์. (2541). การเปรียบเทียบความคลาดเคลื่อนแบบปล่อยคะแนนความคลาดเคลื่อนแบบฮาโล และความคลาดเคลื่อนแบบขัดแย้งของมาตรประมาณค่าที่มีจำนวนช่วงมาตรต่างกัน เมื่อสถานภาพของผู้ประเมินต่างกัน. ปรินูญญานิพนธ์ ค.ม. (การวิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- บุญธรรม กิจปรีดาบริสุทธิ์. (2531). เทคนิคการสร้างเครื่องมือรวบรวมข้อมูลสำหรับทำวิจัย. กรุงเทพฯ : โรงพิมพ์ศรีอนันต์.

- บุญเชิด ภิญโญนนท์พงษ์. (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์. ปรินทิพนิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- (2545). การวัดประเมินการเรียนรู้. ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- บัวทอง สว่างโสภาคกุล. (2529). จิตวิทยาทั่วไป. พิมพ์ครั้งที่ 3, กรุงเทพฯ : ภาควิชาจิตวิทยา มหาวิทยาลัยเกษตรศาสตร์.
- ประภัสสร สุขชื่น. (2539). การศึกษาคุณภาพแบบทดสอบวัดบุคลิกภาพความเชื่อมั่นในตนเองระหว่างรูปแบบมาตราขั้วทางเลือกทางเดียว (UNIPOLAR) และมาตราขั้วทางเลือกสองทาง (BIPOLAR). ปรินทิพนิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- เปรมวดี กาญจนวีระ. (2541). ความสัมพันธ์ระหว่างการทำให้ตนเองเสียเปรียบกับการดำรงแรงจูงใจภายใน ในการทำกิจกรรม. วิทยานิพนธ์ ค.ม. (จิตวิทยาสังคม) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พนิดา ฟุ้งเกียรตินำสุข. (2540). การประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่าที่มีรูปแบบการตอบแตกต่างกัน. ปรินทิพนิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- เพ็ญศรี ลาหล้าเลิศ. (2534). การเปรียบเทียบค่าสถิติของมาตรประมาณค่าที่มีรูปแบบคำตอบต่างกัน ในการประเมินตนเองของนักเรียนชั้นมัธยมศึกษาตอนปลาย. วิทยานิพนธ์ ค.ม. (การวิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. อัดสำเนา.
- พวงรัตน์ ทวีรัตน์. (2543). วิธีการวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์. พิมพ์ครั้งที่ 5 กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ .
- มาลี เตชะไตรศร. (2537). ปฏิสัมพันธ์ของจำนวนช่วงกับประเภทของคำถามตรงกลางที่มีต่อการเลือกตอบคำถามตรงกลางในการประเมินค่า. วิทยานิพนธ์ ค.ม. (การวิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- วิเชียร เกตุสิงห์. (2530). หลักการสร้าง และวิเคราะห์เครื่องมือที่ใช้ในการวิจัย. กรุงเทพฯ : ไทยวัฒนาพานิช.

- วิยะดา หอธรรมอนันต์. (2526). การเปรียบเทียบความเที่ยงแบบสอบถามซ้ำของ
มาตรวัดทัศนคติแบบลิเคอร์ตที่มีจำนวนรายการคำตอบต่างกัน. วิทยานิพนธ์
ค.ม. (การวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
ถ่ายเอกสาร.
- วิรัช วรรณรัตน์. (2532). การตรวจคุณภาพเครื่องมือในการวัดผลและการวิจัย.
สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศิริชัย กาญจนวาสี. (2537). การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย. โรงพิมพ์
จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ศิรินาถ สิงหาแก้ว. (2536). การเปรียบเทียบค่าสถิติของมาตรฐานค่าที่มีจำนวน
ช่วงต่างกัน. วิทยานิพนธ์ ค.ม. (การวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย
จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- สุภาพ วาดเขียน. (2525). เครื่องมือวิจัยทางสังคมศาสตร์ ลักษณะที่ดีชนิดและวิธีหา
คุณภาพ. กรุงเทพฯ : ไทยวัฒนาพานิช.
- สุภาวดี ดั่งบุผา. (2534). การสร้างแบบทดสอบวัดความสามารถในการคิดสร้างสรรค์
ทางคณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 กรุงเทพมหานคร.
ปริญญาณิพนธ์ กศ.ม. (วัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย
ศรีนครินทรวิโรฒ. อัดสำเนา.
- สุมาลี จันทรชลอ. (2542). การวัดและประเมินผล. กรุงเทพฯ : ศูนย์สื่อเสริมกรุงเทพ.
- อนันต์ ศรีโสภ. (2521). หลักการวิจัยเบื้องต้น. กรุงเทพฯ : วัฒนาพานิช.
- , (2525). การวัดผลทางการศึกษา. พิมพ์ครั้งที่ 3. กรุงเทพฯ : วัฒนาพานิช.
- อุทุมพร จามรมาน. (2538). ทฤษฎีการวัดผลทางจิตวิทยา. กรุงเทพฯ : คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย.
- Allen, Mary J and Wendy M. Yen. (1979). *Introduction to Measurement Theory*.
California : Brooks/Cole.
- Anastasi, Anne. (1968). *Psychological Testing*. New York : McGraw-Hill, Inc.
- , (1976). *Psychological Testing*. New York : McGraw-Hill, Inc.
- Aiken, Lewis R. (1983,Sum). "Number of response categories and statistics on
a teacher rating scale," *Education and Psychological Measurement*. 43
(2):397-401.
- Bacon, D.R. (1995,Jun). "composite Reliability in Structural Equations Modeling."
Educational and Psychological Measurement. 55 (3) : 394 – 406.

Bending. A.W. (1953). "The reliability of self – ratings as a function of the amount of Verbal anchoring and of number of categories on the scale. " *Journal of Applied Psychology*. 37 : 38 – 41.

Cunningham.G.k. (1986). *Educational and Psychological Measurement*. New York : Macmillan Publishing.

Ferguson, George. (1966). *A Statistical Analysis in Psychology and Education*. 3rd ed: New York : McGraw-Hill, Book Company.

Finn, R.H. (1972,Sum). "Effect of some Variations in Rating Scale Characteristic on the Means and Reliability of Ratings," *Educational and Psychological Measurement*. 32(2) : 255 – 265.

Good, Carter Vicctor. (1973). *Dictionary of Education*. New York : McGraw – Hill.

Guilford J.P. (1954). *Psychometric Methods*. New York : McGraw – Hill.

Guthrie, Edwin Ray Psychology. (1944). *A First Course in Human Behavior*. New York : Harper.

Lindvall C.M. and Anthony J. Nitko. (1967). *Measuring Pupil Achievement and Aptitude*. New York : Harcourt Brace Jovanovich, Inc.

Lord, Frederic M. and McIvin, R. Novick. (1968). *Statistical Theories of Mental Test Score*. California : Horcourt Brace Jovanovich, Inc.

Mehrens, William A, and Irvin J, Lehmann. (1973). *Measurement and Evaluation in Education and Psychology*. New York : Rinehart and Winston.

----- (1984). *Measurement and Evaluation in Educational and Psychology*. 3rd ed Japan : CBS College Publishing.

Mischel, Walter. (1976). *Introduction Personality*. 2nd ed. New York : Holt Rinehart and Winton.

Newstead, S.E. and Arnold, J. (1989,Spr). "The Effect of Response Format on Rating in Teaching," *Educational and Psychological Measurement*. 49(1) : 33 – 43.

/ Novick. M. R. and C. Lewis. (1967). "Coefficient Alpha and the Reliability of Composite Measurement," *Psychometrika*. 32 : 1 – 13.

Nunnally Jam C. (1964). *Education Measurements and Evolution*. New York : McGraw – hill, Book Company.

- Peter, David L., and McCormic, Ernest J. (1966). Comparative Reliability of Numerically Anchored versus Job – Task Anchored Rating Scales. *Journal of Applied Psychology*. 50(1) : 92 – 96.
- Rummel, J.F. (1964). *An Introduction to Research Procedures in Education*. New York : Haper & Row Publication.
- Standley Julian C. and Kenncth D. Hopkins. (1972). *Education and Psychology and Education*. 5th ed. Engliwoods cliff : Pretice – hall.
- Symonds, P.M. (1925). "Notes on rating," *Journal of General Education*. 105 : 188-196.
- Thorndike, R.L. and Hagen. E.P. 1977. *Measurement and Evaluation in Psychology and Evaluation*. 4th ed. New York : John Wiley and Song.
- Tull, D.S. and Hawkins, D.I. (1993). *Marketing Research*. New York : Macmillan.
- Turney Billy L. and Robb George P. (1971). *Research in Education and Introduction*. Illinois : The Dryden Prss Inc.
- Vockell, Edward L. (1995). *Educational Research*. New York:Macmillan Publishing Co.
- Wells, William D. (1960). "Four semantic Rating Scales Compared." *Journal of Applied Asychology*. 44 : 393 – 397.

ภาคผนวก

ภาคผนวก ก
รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญ

ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช

ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

อาจารย์ชวลิต รวยอาจิณ

ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

อาจารย์วงศ์สันติ รัตนบำรุง

ผู้อำนวยการโรงเรียนบ้านอาโพน จังหวัดสุรินทร์

อาจารย์เรืองอุไร คมรไชย

วิทยาลัยพยาบาลบรมราชชนนี สรรพสิทธิประสงค์

อาจารย์ชนม์ชกรณั ทินบัว

ผู้ช่วยศึกษานิเทศก์สำนักงานการประถมศึกษา กิ่งอำเภอ
สากเหล็ก จังหวัดพิจิตร

ภาคผนวก ข
ผลการวิเคราะห์รายชื่อ

ตาราง 8 แสดงค่าความสอดคล้อง (IOC) ของมาตรฐานประเมินค่าลักษณะนิสัยการทำให้ตนเอง
เสียเปรียบ

ข้อที่	IOC
1	1.00
2	1.00
3	1.00
4	0.80
5	1.00
6	1.00
7	1.00
8	0.80
9	1.00
10	0.80
11	1.00
12	1.00
13	1.00
14	1.00
15	1.00
16	1.00
17	0.80
18	0.80
19	1.00
20	1.00
21	1.00
22	0.60
23	1.00
24	1.00
25	1.00
26	1.00
27	1.00

ตาราง 9 แสดงค่าอำนาจจำแนก (r) ของมาตรประเมินค่าลักษณะนิสัยการทำให้ตนเอง
เสียเปรียบ

ข้อที่	r
1	.995
2	.994
3	.994
4	.995
5	.993
6	.996
7	.997
8	.995
9	.995
10	.995
11	.996
12	.995
13	.997
14	.995
15	.994
16	.996
17	.993
18	.996
19	.996
20	.995
21	.994
22	.991
23	.994
24	.996
25	.995
26	.995
27	.994

ภาคผนวก ค
เครื่องมือที่ใช้ในการวิจัย

มาตรวัดลักษณะนิสัยทำให้ตนเองเสียเปรียบ [แปลจาก Self – handicapping Scale
สร้างโดย Jones และ Rhodewalt (1982)]

ตอนที่ 1

คำชี้แจง โปรดเขียนเครื่องหมาย ✓ ลงใน ☐ หน้าข้อความที่ตรงกับสภาพที่เป็นจริงของนักเรียน

1. เพศ ☐ ชาย ☐ หญิง
2. โรงเรียน
3. ชั้น เลขที่

หมายเหตุ

- ข้อมูลจากแบบสอบถามนี้ เพื่อพัฒนาเครื่องมือมาตรฐานประเมินค่า เพื่อก่อให้เกิดการพัฒนาและความเจริญงอกงามทางการศึกษา ดังนั้น ขอให้นักเรียนทุกคนโปรดเติมตัวเลขลงในช่องว่างหน้าข้อความที่ตรงกับความรู้สึกที่แท้จริง เพื่อให้ได้ข้อมูลที่มีคุณค่าต่อการศึกษามากที่สุด

- ข้อมูลจากแบบสอบถามนี้ ผู้วิจัยรับรองว่าจะไม่มีการนำไปเปิดเผยและไม่มีผลกระทบใดๆ ต่อตัวนักเรียนทุกคนทั้งสิ้น

ขอขอบคุณในความร่วมมือ

ผู้วิจัย

ฉันทลักษณ์ ศรีผา

ข้อกระทง

1. เวลาที่ทำอะไรผิด ข้าพเจ้ามักจะโทษสิ่งแวดล้อมเป็นอันดับแรก
2. ข้าพเจ้ามักทำงาน เมื่อจวนเจียนจะหมดเวลา
3. เวลาจะสอบข้าพเจ้ามักอ่านหนังสือมากเกินไปจนรู้สึกสับสน
4. เวลาจะทำสิ่งใดข้าพเจ้ามักจะเตรียมตัวมากจนรู้สึกเครียด
5. บ่อยครั้ง ข้าพเจ้าอารมณ์ไม่ดีมักเกิดจากเพื่อน ๆ มากระเซ้าเย้าแหย่
6. ข้าพเจ้ามักทำอย่างให้ดีที่สุดทุกครั้ง ไม่ว่าจะทำอะไรก็ตาม
7. ก่อนที่ข้าพเจ้าจะทำกิจกรรมที่สำคัญใดๆ ข้าพเจ้าต้องแน่ใจว่าตนเองมีความพร้อมและมีความสามารถเพียงพอ
8. ก่อนสอบข้าพเจ้ามักจะวิตกกังวลอย่างมาก
9. ข้าพเจ้าไม่ค่อยมีสมาธิอ่านหนังสือเพราะมีคนมารบกวน
10. ข้าพเจ้ามักคิดฟุ้งซ่านทำให้เสียสมาธิเวลาอ่านหนังสือ
11. ข้าพเจ้ามักไม่คาดหวังในกิจกรรมที่มีการแข่งขันสูงเพราะกลัวผิดหวัง
12. ข้าพเจ้าอยากให้คนอื่นชื่นชมความตั้งใจในการทำงานมากกว่าผลของงาน
13. ถ้าข้าพเจ้าพยายามมากขึ้น คงทำได้ดีกว่านี้
14. ข้าพเจ้าชอบความสุขเล็กๆ น้อยๆ ในปัจจุบันมากกว่าความสุขมากมายในอนาคตที่ยังมาไม่ถึง
15. ข้าพเจ้าไม่ชอบตกอยู่ในสถานการณ์ที่ทำงานล้มเหลว
16. สักวันข้าพเจ้าอาจจะได้ทุกสิ่งตามที่ต้องการ
17. บางครั้งข้าพเจ้าชอบที่จะเจ็บป่วยเล็กน้อย เพราะมันช่วยลดปล่อยความกดดันที่มี
18. ข้าพเจ้าจะทำได้ดีกว่านี้ ถ้าไม่ปล่อยให้อารมณ์เข้ามาเกี่ยวข้อง
19. เวลาทำอะไรล้มเหลว ข้าพเจ้ามักจะปลอบใจตนเองโดยนึกถึงความสำเร็จที่ผ่านมา
20. ข้าพเจ้ามักจะหาเหตุผลเข้าข้างตนเองเสมอ เมื่อทำไม่ได้ตามความคาดหวังของคนอื่น
21. ข้าพเจ้ามักทำอะไรไม่ค่อยสำเร็จเพราะดวงไม่ค่อยดี
22. ข้าพเจ้าไม่เคยใช้ยาที่ขัดขวางความสามารถในการทำกิจกรรมใดๆ
23. ข้าพเจ้ามักตามใจตนเองในการรับประทานอาหารและเครื่องดื่ม
24. ก่อนที่จะทำกิจกรรมที่สำคัญใดๆ ข้าพเจ้าจะพักผ่อนให้เพียงพอ
25. ข้าพเจ้าไม่เคยให้ปัญหาทางอารมณ์มีอิทธิพลในการดำเนินชีวิต
26. เมื่อใดที่ข้าพเจ้ากังวลเกี่ยวกับงานที่กำลังทำ งานนั้นมักจะสำเร็จลงด้วยดี
27. บางเวลาเมื่อข้าพเจ้ารู้สึกหดหู่ แม้แต่งานง่ายๆ ก็กลายเป็นยาก

ในตอนที 2 ผู้วิจัยได้ดัดแปลงข้อกระทงความลักษณะนิสัยการทำให้ตนเองเสียเปรียบดั่งกล่าวข้างต้น ให้เป็นมาตรประเมินค่าที่มีรูปแบบ ต่าง ๆ ทำให้ได้เครื่องมือทั้งหมด 18 ฉบับดังนี้

- ฉบับที่ 1 แบบเติมตัวเลข ที่มีจำนวนมาตรา 4 มาตรา
- ฉบับที่ 2 แบบเติมตัวเลข ที่มีจำนวนมาตรา 5 มาตรา
- ฉบับที่ 3 แบบเติมตัวเลข ที่มีจำนวนมาตรา 6 มาตรา
- ฉบับที่ 4 แบบเติมตัวเลข ที่มีจำนวนมาตรา 7 มาตรา
- ฉบับที่ 5 แบบเติมตัวเลข ที่มีจำนวนมาตรา 8 มาตรา
- ฉบับที่ 6 แบบเติมตัวเลข ที่มีจำนวนมาตรา 9 มาตรา
- ฉบับที่ 7 แบบบรรยาย ที่มีจำนวนมาตรา 4 มาตรา
- ฉบับที่ 8 แบบบรรยาย ที่มีจำนวนมาตรา 5 มาตรา
- ฉบับที่ 9 แบบบรรยาย ที่มีจำนวนมาตรา 6 มาตรา
- ฉบับที่ 10 แบบบรรยาย ที่มีจำนวนมาตรา 7 มาตรา
- ฉบับที่ 11 แบบบรรยาย ที่มีจำนวนมาตรา 8 มาตรา
- ฉบับที่ 12 แบบบรรยาย ที่มีจำนวนมาตรา 9 มาตรา
- ฉบับที่ 13 แบบกราฟฟิก ที่มีจำนวนมาตรา 4 มาตรา
- ฉบับที่ 14 แบบกราฟฟิก ที่มีจำนวนมาตรา 5 มาตรา
- ฉบับที่ 15 แบบกราฟฟิก ที่มีจำนวนมาตรา 6 มาตรา
- ฉบับที่ 16 แบบกราฟฟิก ที่มีจำนวนมาตรา 7 มาตรา
- ฉบับที่ 17 แบบกราฟฟิก ที่มีจำนวนมาตรา 8 มาตรา
- ฉบับที่ 18 แบบกราฟฟิก ที่มีจำนวนมาตรา 9 มาตรา

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ-นามสกุล	นางสาวฉันทลักษณ์ ศรีผา
วันเดือนปีเกิด	3 กันยายน 2519
สถานที่เกิด	อำเภอโพธาราม จังหวัดร้อยเอ็ด
สถานที่อยู่ปัจจุบัน	100 หมู่ 7 ตำบลโพธาราม อำเภอโพธาราม จังหวัดร้อยเอ็ด 45240
ตำแหน่งหน้าที่การงานในปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ประวัติการศึกษา	
พ.ศ. 2535	ชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสตรีศึกษา
พ.ศ. 2538	ชั้นมัธยมศึกษาปีที่ 6 โรงเรียนร้อยเอ็ดวิทยาลัย
พ.ศ. 2543	วท. บ. (คณิตศาสตร์) มหาวิทยาลัยศรีนครินทรวิโรฒ
พ.ศ. 2545	กศ. ม. (การวิจัยและสถิติทางการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ