

การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทาง
การเรียนรู้ตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ
ด้วยวิธีการตรวจสอบต่างกัน

ปริญญาณิพนธ์
ของ
ชุติมา แสงदारารัตน์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
พฤษภาคม 2546
ลิขสิทธิ์เป็นของ มหาวิทยาลัยศรีนครินทรวิโรฒ

การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทาง
การเรียนรู้ตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ
ด้วยวิธีการตรวจสอบต่างกัน

บทคัดย่อ
ของ
ชุตินา แสงดารารัตน์

- 3 ก.ย. 2546

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
พฤษภาคม 2546

๒ 227065

ชุตินา แสงดารารัตน์. (2545). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบด้วยวิธีการตรวจสอบต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา).

กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ , อาจารย์ชวลิต รวยอาจิน.

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อเปรียบเทียบผลการตรวจสอบการทำหน้าที่ของแบบทดสอบวัดความถนัดทางการเรียน ตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบด้วยวิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน โดยเปรียบเทียบจำนวนข้อที่ทำหน้าที่ต่างกัน และค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดข้อสอบที่ทำหน้าที่ต่างกันออกแล้ว ในการศึกษาครั้งนี้มีกลุ่มตัวอย่างเป็น นักเรียนชั้นประถมศึกษาชั้นปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ จำนวน 585 คน ได้จากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) เครื่องมือที่ใช้ในการศึกษาเป็น แบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 90 ข้อ ผลการวิจัยพบว่า

1. จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผลด้วย วิธีการตรวจสอบต่างกัน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกัน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ ได้จำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านภาษา และด้าน ตัวเลข ด้วยวิธีการตรวจสอบต่างกัน บ่งชี้ข้อที่ทำหน้าที่ต่างกัน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้าน ตัวเลข หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกโดยใช้วิธีการตรวจสอบต่างกัน พบว่า

2.1 แบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2.2 แบบทดสอบวัดความถนัดทางการเรียน ด้านภาษา จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จากกลุ่มอ้างอิงที่เป็นความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2.3 แบบทดสอบวัดความถนัดทางการเรียน ด้านตัวเลข จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

A COMPARISON OF SCHOLASTIC APTITUDE TEST OF DIFFERENTIAL ITEM
FUNCTIONING DETECTION BY FAMILIARITY INTEREST AND PLEASANTNESS
IN TESTS USING BY DIFFERENT DETECTING METHODS

AN ABSTRACT
BY
CHUTIMA SAENGDARARAT

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University
May 2003

Chutima Saengdararat. (2002). *A Comparison of Scholastic Aptitude Test of Differential Item Functioning Detection by Familiarity Interest and Pleasantness in Tests using by Different Detecting Methods*. Master Thesis, M.Ed. (Education Measurement). Bangkok : Graduate School , Srinakharinwirot University. Advisor Committee : Assoc. Prof Dr.Boonchird Pinyoanuntapong, Mr. Chawalit Ruayajin.

This study aimed to compare scholastic aptitude test of differential item functioning detection according to familiarity, interest, and preference by Mantel-Haenszel method and the analysis of variance method through comparing number of differential functioning item and reliability of test after screening. The subjects consisted of 585 Pratom Suksa VI students in schools under the Bangkok Metropolitan during 2nd semester of 2000 academic year, through stratified random sampling. The instrument for the study included reasoning, verbal, and numerical scholastic aptitude test with 90 items of five multiple choices.

The results revealed that :

1. There was a statistical significant difference at .05 level among number of differential functioning item of reasoning scholastic aptitude test through different detection indicating differential functioning item according to criterion group of gender and interest, whereas there was no significant difference among those of criterion group of familiarity and pleasantness.

There was no statistical significant difference among number of differential functioning item of verbal and numerical scholastic aptitude test through different detection indicating differential detection according to criterion group of gender, familiarity, interest, and pleasantness.

2. The reliability of reasoning, verbal, and numerical scholastic aptitude test after screening showed that ;

- 2.1 There was a statistical significant difference at .05 level among the reliability of reasoning scholastic aptitude test according to criterion group of gender and interest, whereas there was no statistical significant difference among those according to criterion group of familiarity and pleasantness.

- 2.2 There was no statistical significant difference among the reliability of verbal scholastic aptitude test according to criterion group of gender, familiarity, and pleasantness, whereas there was a statistical significant difference at .05 level among those according to criterion group of interest.

- 2.3 There was no statistical significant difference among the reliability of numerical scholastic aptitude test according to criterion group of gender, familiarity, interest, and pleasantness.

ปริญญานิพนธ์

เรื่อง

การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทาง
การเรียนรู้ ตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ
ด้วยวิธีการตรวจสอบต่างกัน

ของ

นางสาวชุติมา แสงदारารัตน์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร.นภาพร หะวานนท์)

วันที่...๑...เดือน พฤษภาคม พ.ศ. 2546

คณะกรรมการสอบปริญญานิพนธ์

.....ประธาน

(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์)

.....กรรมการ

(อาจารย์ชวลิต รวยอาจิณ)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(อาจารย์ ดร.สุพร เข้มเฮง)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(อาจารย์ ดร.ละเอียด รักษ์เฝ้า)

ปริญญาโทฉบับนี้ได้รับ
ทุนอุดหนุนการวิจัย จาก
มหาวิทยาลัยศรีนครินทรวิโรฒ ประจำปีการศึกษา 2544
ผู้วิจัยขอขอบคุณมา ณ ที่นี้

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยความอนุเคราะห์และความกรุณาอย่างสูงจาก
รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์ และอาจารย์ชวลิต รวยอาจิน ประธานและกรรมการ
ที่ปรึกษาปริญญานิพนธ์ ดร.สุพร เข้มเฮง และ ดร.ละเอียด รักษาเผ่า กรรมการสอบปริญญานิพนธ์ ที่กรุณา
ให้คำปรึกษา แนะนำ ชี้แนะแนวทาง ตรวจ แก้ไขข้อบกพร่องต่างๆ จนปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไป
ด้วยดี ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างสูง

กราบขอบพระคุณคณาจารย์ ในภาควิชาการวัดผลและวิจัยการศึกษาทุกท่าน ที่ได้อบรมสั่งสอน
วิชาความรู้ในศาสตร์ที่เกี่ยวข้องกับการวัดผลการศึกษา และได้กรุณาให้คำแนะนำ แก้ไขข้อบกพร่องต่างๆ ใน
การสร้างเครื่องมือครั้งนี้

ขอขอบคุณผู้บริหาร คณะครู-อาจารย์ทุกท่านของโรงเรียนที่เป็นกลุ่มตัวอย่าง และขอขอบใจ
นักเรียนทุกคนที่ให้ความร่วมมือในการเก็บข้อมูลเป็นอย่างดี

ขอขอบคุณคณะผู้บริหาร คณาจารย์ และพนักงาน วิทยาลัยเซนต์หลุยส์ ที่ให้โอกาส และ
เป็นกำลังใจแก่ผู้วิจัยในการทำปริญญานิพนธ์ฉบับนี้

ขอขอบคุณ พี่ๆ เพื่อนๆ และน้องๆ นิสิตปริญญาโท วิชาเอกการวัดผลการศึกษา ภาคปกติ แผน ก
รุ่นที่ 34 และ แผน ข รุ่นที่ 2 รวมทั้งกัลยาณมิตรทุกท่านที่ช่วยเหลือ และเป็นกำลังใจในการทำปริญญานิพนธ์
ฉบับนี้

สุดท้ายนี้ผู้วิจัยขอกราบขอบพระคุณ คุณพ่ออนันต์ – คุณแม่ยุวดี แสงदारารัตน์ ที่ เป็นกำลังใจ
และคอยดูแล ช่วยเหลือทุกสิ่งทุกอย่างแก่ผู้วิจัยอย่างดีที่สุดเสมอมา และขอขอบคุณ ญาติ พี่น้อง กัลยาณมิตร และ
ผู้ที่มีส่วนเกี่ยวข้องทุกท่าน ที่สนับสนุนทั้งกำลังใจ กำลังใจ และดูแลช่วยเหลือในทุกสิ่งทุกอย่างแก่ผู้วิจัย
อย่างดีที่สุดเสมอมา

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาแด่ บิดา
มารดา และครูอาจารย์ ตลอดจนผู้มีพระคุณทุกท่านที่ได้ให้การอบรมสั่งสอน จนผู้วิจัยประสบความสำเร็จใน
การศึกษา

ชุตินา แสงदारารัตน์

สารบัญ

บทที่		หน้า
1	บทนำ.....	1
	ภูมิหลัง.....	1
	ความมุ่งหมายของการศึกษาค้นคว้า.....	3
	ความสำคัญของการศึกษาค้นคว้า.....	3
	ขอบเขตของการศึกษาค้นคว้า.....	3
	ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย.....	3
	ตัวแปรที่ศึกษา.....	4
	นิยามศัพท์เฉพาะ.....	4
	สมมติฐานของการศึกษาค้นคว้า.....	5
2	เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
	เอกสารที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ.....	6
	เอกสารที่เกี่ยวกับความหมายและทฤษฎีที่เกี่ยวข้องกับความถนัด.....	12
	งานวิจัยที่เกี่ยวข้อง.....	21
	งานวิจัยในต่างประเทศ.....	21
	งานวิจัยในประเทศ.....	25
3	วิธีดำเนินการศึกษาค้นคว้า.....	30
	การกำหนดประชากรและเลือกกลุ่มตัวอย่าง.....	30
	การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	31
	ลักษณะเนื้อหาของแบบทดสอบ.....	34
	การจัดกระทำและการวิเคราะห์ข้อมูล.....	39
	ขั้นตอนในการวิเคราะห์ข้อมูล.....	39
	สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	39
4	ผลการวิเคราะห์ข้อมูล.....	44
	สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	44
	การนำเสนอผลการวิเคราะห์ข้อมูล.....	45
	ผลการวิเคราะห์ข้อมูล.....	45

สารบัญ (ต่อ)

บทที่	หน้า
5 สรุปผล อภิปราย และข้อเสนอแนะ.....	55
จุดมุ่งหมายของการศึกษาค้นคว้า.....	55
กลุ่มตัวอย่าง.....	55
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	55
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	55
การเสนอผลการวิเคราะห์ข้อมูล.....	56
สรุปผลการวิเคราะห์ข้อมูล.....	56
อภิปรายผล.....	57
ข้อเสนอแนะ.....	58
บรรณานุกรม.....	59
ภาคผนวก.....	64
ภาคผนวก ก.....	65
ภาคผนวก ข.....	67
ภาคผนวก ค.....	69
ภาคผนวก ง.....	71
ประวัติย่อผู้วิจัย.....	74

บัญชีตาราง

ตาราง		หน้า
1	แสดงความถี่ของผู้สอบที่ตอบข้อสอบถูกและผิดในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบสำหรับช่วงคะแนน j	8
2	แสดงสัดส่วนการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่คะแนน j	9
3	จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามขนาดโรงเรียน.....	31
4	จำนวนนักเรียนกลุ่มอ้างอิง กลุ่มเปรียบเทียบของแบบทดสอบวัดความถนัดทางการเรียน.....	46
5	การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้วยวิธีการตรวจสอบต่างกัน.....	47
6	การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียน ด้านภาษา ด้วยวิธีการตรวจสอบต่างกัน.....	48
7	การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียน ด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน.....	49
8	การเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน.....	50
9	ค่าความเชื่อมั่นของแบบทดสอบหลังการปรับขยายให้แบบทดสอบมีจำนวนข้อเท่ากับ 30 ข้อ.....	52
10	การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดข้อสอบที่ทำหน้าที่ต่างกัน ออกแล้วด้วยด้วย UX_j	54
11	ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดความถนัดทางการเรียน.....	68
12	จำนวนกลุ่มอ้างอิง กลุ่มเปรียบเทียบ ของแบบทดสอบวัดความถนัดทางการเรียน.....	70
13	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล.....	72
14	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านภาษา.....	73
15	การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านตัวเลข.....	74

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1. โครงสร้างความสามารถของทฤษฎีหลายองค์ประกอบ.....	14
2. โครงสร้างความสามารถทางสมองของทฤษฎีไฮราคัล.....	16
3. โครงสร้างความสามารถทางสมองของทฤษฎีโครงสร้างสามมิติของปัญญา.....	18
4. โครงสร้างทฤษฎีเขาวงกตปัญญาของสเตอร์นเบิร์ก.....	19
5. ลำดับขั้นตอนการสร้างเครื่องมือ.....	32

บทที่ 1

บทนำ

ภูมิหลัง

การทดสอบเป็นการค้นและพัฒนาสมรรถภาพของมนุษย์ ซึ่งมนุษย์แต่ละคนมีความสามารถที่จะเรียนรู้ได้แตกต่างกัน เพื่อพัฒนาและกระตุ้นให้ผู้เรียนเกิดการเรียนรู้ และตรวจดูความรู้ความสามารถของผู้เรียน ชวาล แพรัตกุล(2517:50) กล่าวว่า ความถนัดเป็นขีดความสามารถทางสมองของบุคคลที่มีต่อการเรียนรู้หรือทำกิจกรรมต่างๆ เมื่อได้รับประสบการณ์และการฝึกที่เหมาะสม เช่นเดียวกับ แซค (Sax.1989:338-357) ที่กล่าวว่าแบบทดสอบวัดความถนัด เป็นการวัดสติปัญญา หรือวัดความสามารถทางสมองที่ถูกออกแบบขึ้นมาเพื่อพยากรณ์ว่าแต่ละบุคคลจะประสบความสำเร็จได้ดีเพียงใด ก่อนที่จะได้รับการฝึกหัด การคัดเลือก หรือการจัดตำแหน่ง ซึ่งจากความหมายที่กล่าวมานี้เป็นความหมายของความถนัดที่เป็นแนวโน้มที่จะนำไปสู่ความสำเร็จในกิจกรรมที่บุคคลนั้นๆ มีความถนัด จึงมีส่วนช่วยให้ผู้สอนทราบว่าบุคคลนั้นจะสามารถเรียนได้ดีเพียงใด มีโอกาสที่จะเพิ่มความสำเร็จให้กับตนเองหรือไม่ และมีความสามารถผ่านเกณฑ์ที่ต้องการ ในการวัดผลและการประเมินผลทางการศึกษานั้น มีแบบทดสอบปรนัยเป็นเครื่องมือที่นิยมใช้กันมาก และแบบทดสอบที่นำมาใช้ในการวัดจะต้องเป็นแบบทดสอบที่ดีมีคุณภาพเพื่อให้ผลการวัดที่เชื่อถือได้ มีประสิทธิภาพ สามารถพยากรณ์หรืออธิบายความสามารถของผู้สอบได้ คุณสมบัติของแบบทดสอบที่ดีประกอบด้วย 1.ความเที่ยงตรง (Validity) 2.ความเชื่อมั่น (Reliability) 3.ความยากง่าย (Difficulty) ที่พอเหมาะ 4.ความเป็นปรนัย (Objectivity) 5.ความยุติธรรม (Fairness) 6.ความจำเพาะเจาะจง (Specificity) และมีค่าอำนาจจำแนก (Discrimination) จากคุณลักษณะของข้อสอบข้างต้นจะพบว่าความยุติธรรม เป็นคุณลักษณะที่สำคัญอย่างหนึ่งของข้อสอบ ข้อสอบที่ขาดความยุติธรรม หมายถึงข้อสอบที่เปิดโอกาสให้ผู้สอบกลุ่มใดกลุ่มหนึ่งในระดับความสามารถเท่ากัน มีโอกาสที่จะตอบถูกต่างกัน อาจมีความแตกต่างกันในด้าน ภาษา เพศ ศาสนา เชื้อชาติ วัฒนธรรม ภูมิสำเนา หรือ ระดับสติปัญญา ซึ่งข้อสอบในลักษณะนี้เรียกว่า “ ข้อคำถามลำเอียง หรือ ข้อสอบลำเอียง (Item Bias) ” เป็นคำที่ใช้กล่าวถึงตัวข้อสอบและให้ความหมายที่คลุมเครือในการใช้เกณฑ์ตัดสินเรื่องความลำเอียง ซึ่งต่อมาในระยะหลัง นักวิจัย และนักวัดผลได้เปลี่ยนมาใช้คำว่า “ การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Function หรือ DIF) ” แทน (Holland&Wainer.1993) เป็นวิธีการที่ใช้เทคนิควิธีการทางสถิติเข้ามาตรวจสอบ ความแตกต่างกันของคุณสมบัติทางสถิติของข้อสอบในกลุ่มผู้สอบที่ต่างกลุ่มกัน หลังจากควบคุมความแตกต่างกันในความสามารถของกลุ่ม (สมศักดิ์ จันผ่อง. 2541: 89) คือ เกิดความลำเอียงหรือความไม่ยุติธรรมขึ้นในการสอบ เพราะหากข้อสอบหรือแบบทดสอบมีความลำเอียงโดยเข้าข้างกลุ่มใดกลุ่มหนึ่ง หรือเอื้อประโยชน์ต่อคนบางกลุ่มแล้ว จะทำให้ผลการสอบขาดความน่าเชื่อถือ และเกิดความไม่เสมอภาคในการตัดสิน

กระบวนการในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสามารถทำได้ 2 ขั้นตอน ดังนี้

1. การตรวจสอบในขั้นตอนการสร้างข้อสอบ โดยให้ผู้เชี่ยวชาญพิจารณาว่า ข้อสอบวัดได้ตรงจุดมุ่งหมายหรือพฤติกรรมและเนื้อหาที่ต้องการวัดหรือไม่จากตารางวิเคราะห์หลักสูตรหรือเป็นการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง และความเที่ยงตรงเชิงเนื้อหา
2. การตรวจสอบโดยใช้สถิติ หลังจากการที่นำแบบทดสอบไปใช้แล้วสามารถตรวจสอบได้ 2 แนวทาง
 - 2.1 ใช้วิธีภายนอก (External Method) หรือเกณฑ์ภายนอก(External Criterion) โดยใช้

แบบทดสอบที่สร้างขึ้นหาความสัมพันธ์กับเกณฑ์มาตรฐาน เช่น เกณฑ์เฉลี่ย ผลการปฏิบัติงาน แต่วิธีนี้ไม่เป็นที่นิยม เพราะเกณฑ์ภายนอกที่นำมาหาความสัมพันธ์ ถ้าไม่มีมาตรฐานแล้วจะทำให้การตรวจสอบหาความถูกต้อง

2.2 ใช้วิธีภายใน (Internal Method) หรือเกณฑ์ภายใน (Internal Criterion) เป็นวิธีที่ได้รับความนิยมอย่างมากเพราะสามารถศึกษาจากการตรวจสอบโครงสร้างภายในของข้อสอบ โดยพิจารณาคะแนนที่ได้จากผลตอบของผู้เข้าสอบแต่ละกลุ่มว่าวัดในคุณลักษณะที่ต้องการวัดตามโครงสร้างเช่นเดียวกันไม่ ถือเป็นกระบวนการของการสนับสนุนการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง สามารถแบ่งออกเป็น 3 กลุ่ม ใหญ่ๆ ดังนี้ (Camilli and Shepard.1994:22-131)

2.2.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory) ได้แก่ วิธีการวิเคราะห์ความแปรปรวน (Analysis of Variance) วิธีแปลงค่าความยาก (Transformation Item Difficulty) วิธี Golden Rule วิธีพอยต์ไบซีเรียล (Piont Biserial) ข้อดีของวิธีการในกลุ่มนี้คือ ไม่ยุ่งยาก ประหยัดค่าใช้จ่าย ใช้ตรวจสอบกับกลุ่มตัวอย่างเล็กๆ ได้ ข้อเสีย คือ ค่าสถิติจะเปลี่ยนแปลงไปตามกลุ่มตัวอย่าง ทำให้เกิดความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันที่ได้

2.2.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้ตารางการถ่วง (Contingency Table) การตรวจสอบในแนวนนี้ได้แก่ วิธีแมนเทล-แฮนส์เซล (Mantal-Haenszel) และวิธีไคสแควร์ (Chi-Square Methods) ใช้คะแนนของแบบทดสอบ เป็นเกณฑ์การจับคู่กลุ่มผู้สอบย่อยสองกลุ่มที่ศึกษา ข้อดี ของวิธีการในกลุ่มนี้ คือ กระบวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่ยุ่งยาก ค่าใช้จ่ายในการดำเนินการไม่สูง ใช้กับกลุ่มตัวอย่างขนาดไม่ใหญ่นัก มีหลักการที่ดีในการจับคู่กลุ่มผู้สอบย่อยตามความสามารถของผู้สอบ และมีการทดสอบนัยสำคัญ

2.2.3 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามแนวคิดทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โดยใช้การเปรียบเทียบโค้งลักษณะข้อสอบ (Item Characteristic Curves : ICC) ของกลุ่มผู้สอบย่อยที่ศึกษาตามระดับความสามารถของผู้สอบ ข้อดีของวิธีการกลุ่มนี้ คือ การแก้ไขข้อบกพร่องของทฤษฎีแบบดั้งเดิม การประมาณค่าความสามารถของผู้สอบเป็นอิสระจากค่าความยากของแบบทดสอบข้อเสีย คือ วิธีที่ใช้ซับซ้อน เสียค่าใช้จ่ายในการดำเนินการสูง และใช้กลุ่มตัวอย่างขนาดใหญ่

จะเห็นว่า วิธีการทางสถิติที่ใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มีหลายวิธี วิธีการที่เหมาะสมนั้น แฮมเบิลตัน (Hambleton.1993) ให้ความเห็นว่า ต้องใช้หลักการเปรียบเทียบผลการตอบข้อสอบของกลุ่มผู้สอบอ้างอิง กับกลุ่มเปรียบเทียบ หลังจากการจับคู่กลุ่มผู้สอบ 2 กลุ่ม ตามความสามารถผู้สอบแล้ว ในปัจจุบัน การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทล-แฮนส์เซล เป็นวิธีการที่ได้รับการยอมรับ และเป็นที่ยอมรับจาก สำนักบริการการทดสอบทางการศึกษาแห่งสหรัฐอเมริกา (Education Testing Service : ETS) เนื่องจากวิธีแมนเทล-แฮนส์เซลมีกระบวนการตรวจสอบที่ไม่ยุ่งยากซับซ้อน มีเทคนิคทางสถิติที่เชื่อถือได้ มีการทดสอบนัยสำคัญทางสถิติและขนาดอิทธิพล และสามารถใช้ได้ดีเมื่อกลุ่มตัวอย่างมีขนาดเล็ก เสียค่าใช้จ่ายในการดำเนินการน้อย และผลการวิเคราะห์การทำหน้าที่ต่างกันโดยวิธีแมนเทล-แฮนส์เซล สอดคล้องกับวิธีการวิเคราะห์อื่นๆ เช่น วิธีทฤษฎีการตอบสนองข้อสอบ (Holland and Thayer.1988 ; Millsap and Everson.1993)

ดังนั้น ผู้วิจัยจึงมีความสนใจที่จะศึกษาผลการตรวจสอบข้อสอบที่บ่งชี้การทำหน้าที่ต่างกันระหว่างวิธีแมนเทล-แฮนส์เซล และวิธีวิเคราะห์ความแปรปรวน โดยศึกษาตามความรู้สึกลับเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบวัดความถนัดทางการเรียน จะได้ผลที่แตกต่างกัน

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาในครั้งนี้มีจุดมุ่งหมายสำคัญ เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา ด้านตัวเลข ตามความรู้สึกลึกซึ้ง ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ ด้วยวิธีการตรวจสอบต่างกัน 2 วิธี คือ วิธี แมนเทิล-แฮนด์เชล และ วิธีวิเคราะห์ความแปรปรวน เป็นจุดมุ่งหมายเฉพาะของการศึกษาค้นคว้า ต่อไปนี้

1. เพื่อศึกษาจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน
2. เปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา ด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน
3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา ด้านตัวเลข หลังคัดเลือกข้อที่ทำหน้าที่ต่างกันออก ด้วยวิธีการตรวจสอบต่างกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าในครั้งนี้ จะทำให้ทราบว่าวิธีตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ใช้ตัวแปรในการตรวจสอบการทำหน้าที่ต่างกันที่เป็นเพศ ความรู้สึกลึกซึ้ง ความสนใจ และความพอใจในการทำข้อสอบ เป็นกลุ่มอ้างอิง วิธีใดจะพบจำนวนข้อสอบที่ทำหน้าที่ต่างกันมากที่สุด และทราบค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกแล้ว เพื่อเป็นแนวทางในการศึกษาค้นคว้าวิจัยเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบต่อไป

ขอบเขตของการศึกษาค้นคว้า

ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ จำนวนโรงเรียน 53 โรงเรียน จำนวน 103 ห้องเรียน รวมนักเรียนทั้งสิ้น 2,855 คน จำแนกเป็นนักเรียนชาย 1,491 คน นักเรียนหญิง 1,364 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ ซึ่งได้มาโดยการสุ่มแบบ 2 ขั้นตอน (Two-Stage Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) โรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) จำนวนโรงเรียน 9 โรงเรียน จำนวน 22 ห้องเรียน รวมนักเรียนทั้งสิ้น 585 คน จำแนกเป็นนักเรียนชาย 310 คน นักเรียนหญิง 275 คน

ตัวแปรที่ศึกษา

1. ตัวแปรอิสระ ได้แก่
 - 1.1 วิธีตรวจสอบการทำหน้าที่ต่างกัน ได้แก่
 - 1.1.1 วิธีแมนเทิล-แฮนด์เชล
 - 1.1.2 วิธีวิเคราะห์ความแปรปรวน
2. ตัวแปรตาม ได้แก่
 - 2.1 จำนวนข้อสอบที่ทำหน้าที่ต่างกัน
 - 2.2 ความเชื่อมั่นของแบบทดสอบที่คัดเลือกข้อที่ทำหน้าที่ต่างกันออกไป
3. ตัวแปรที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ
 - 3.1 เพศ
 - 3.2 ความรู้สึกคุ้นเคยในข้อสอบ
 - 3.3 ความรู้สึกสนใจในข้อสอบ
 - 3.4 ความรู้สึกพอใจในข้อสอบ

นิยามศัพท์เฉพาะ

1. ข้อสอบที่ทำหน้าที่ต่างกัน หมายถึง ข้อสอบที่ทำให้ผู้สอบที่มีความสามารถในสิ่งที่ต้องการวัดเท่ากันมีโอกาสตอบข้อสอบข้อนั้นได้ถูกต้องไม่เท่ากัน
2. กลุ่มอ้างอิง หมายถึง กลุ่มผู้สอบที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบเมื่อข้อสอบที่ทำหน้าที่ต่างกันโดยมีโอกาสดังกล่าวได้ถูกต้องมากกว่าผู้สอบอีกกลุ่มหนึ่งซึ่งมีความสามารถเท่ากัน ในการวิจัยครั้งนี้ใช้ กลุ่มเพศชาย กลุ่มที่คุ้นเคย กลุ่มที่สนใจ และกลุ่มที่พอใจในข้อสอบ
3. กลุ่มเปรียบเทียบ หมายถึง กลุ่มผู้สอบที่คาดว่าจะเสียเปรียบจากการตอบข้อสอบที่ทำหน้าที่ต่างกันโดยมีโอกาสดังกล่าวได้ถูกต้องน้อยกว่าผู้สอบอีกกลุ่มหนึ่งซึ่งมีความสามารถเท่ากัน ในการวิจัยครั้งนี้ใช้ กลุ่มเพศหญิง กลุ่มที่ไม่คุ้นเคย กลุ่มที่ไม่สนใจ และกลุ่มที่ไม่พอใจในข้อสอบ
4. วิธีการตรวจสอบการทำหน้าที่ต่างกัน หมายถึง การวิเคราะห์หาดัชนีการทำหน้าที่ต่างกันของข้อสอบ ซึ่งในการวิจัยครั้งนี้ใช้ 2 วิธี คือ
 - 4.1 วิธีแมนเทิล-แฮนด์เชล หมายถึง วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ซึ่งพิจารณาจากความแตกต่างของอัตราส่วนการตอบข้อสอบระหว่างผู้สอบที่มีความสามารถระดับเดียวกันที่พัฒนาโดยฮอลแลนด์และเทเยอร์ (Holland and Thayer.1988) เกณฑ์การตัดสินใจข้อสอบที่ทำหน้าที่ต่างกันตามวิธีแมนเทิล-แฮนด์เชล หมายถึง ข้อสอบที่มีค่า α_{MH} แตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05
 - 4.2 วิธีวิเคราะห์ความแปรปรวน หมายถึง วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบซึ่งพิจารณาจากการวิเคราะห์การแจกแจงระหว่างข้อสอบกับกลุ่มผู้สอบ โดยการทดสอบนัยสำคัญทางสถิติ ที่ระดับความเชื่อมั่นร้อยละ 95 ของค่าสถิติ (F-Value) ถ้าพบว่ามีนัยสำคัญก็จะทำการทดสอบ ความแตกต่างของค่าเฉลี่ย เป็นรายข้อ โดยใช้ t-test ถ้าค่าที่คำนวณได้เท่ากับหรือมากกว่า ค่าวิกฤตของ t ในตาราง (ค่าวิกฤตของ $t_{.05}$ เท่ากับ 1.96) แสดงว่าข้อสอบนั้นมีการทำหน้าที่ต่างกันของข้อสอบ
5. แบบทดสอบวัดความถนัดทางการเรียน หมายถึง ชุดของข้อคำถามที่ใช้วัดความสามารถทางสมองของบุคคลแต่ละบุคคล เพื่อพยากรณ์ว่าจะประสบความสำเร็จในการเรียนเพียงใด ซึ่งสร้างโดยยึดทฤษฎีหลายองค์ประกอบ (Multiple Factor Theory) สำหรับการวิจัยครั้งนี้ ได้สร้างแบบทดสอบดังต่อไปนี้

5.1 แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล (Reasoning Ability) เป็นแบบทดสอบวัดความสามารถในการคิดหาเหตุผลเกี่ยวกับการหาค่า รูปภาพที่ไม่เข้าพวก และการสรุปความ

5.2 แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา (Verbal Ability) เป็นแบบทดสอบวัดความสามารถทางภาษาเกี่ยวกับการหาความหมาย คำตรงข้าม ศัพท์สัมพันธ์ และความเข้าใจภาษา

5.3 แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข (Numerical Ability) เป็นแบบทดสอบวัดความสามารถทางการคิด คำนวณเกี่ยวกับอนุกรมตัวเลข และคณิตศาสตร์เหตุผล

6. ความรู้สึกคุ้นเคยในข้อสอบ (Familiarity) หมายถึง การที่นักเรียนได้เคยทำ หรือเคยได้รับการฝึกทำโจทย์ในลักษณะเช่นนี้มาก่อน ในการวิจัยครั้งนี้ผู้วิจัยกำหนดเป็นความรู้สึกคุ้นเคย และไม่คุ้นเคยในข้อสอบ

7. ความรู้สึกสนใจในข้อสอบ (Interested) หมายถึง การที่นักเรียนมีความรู้สึกนิยม ชมชอบ อยากที่จะทำข้อสอบ ในการวิจัยครั้งนี้ผู้วิจัยกำหนดเป็นความรู้สึกสนใจ และไม่สนใจในข้อสอบ

8. ความรู้สึกพอใจในข้อสอบ (Pleasantness) หมายถึง การที่นักเรียนมีความพร้อมที่จะทำ ยินดีที่ทำข้อสอบ ในการวิจัยครั้งนี้ผู้วิจัยกำหนดเป็นความรู้สึกพอใจ และไม่พอใจในข้อสอบ

9. ผู้เชี่ยวชาญ หมายถึง ผู้มีคุณวุฒิตั้งแต่ระดับปริญญาโทขึ้นไป และมีประสบการณ์ด้านการวัดผลการศึกษา มาไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน

สมมติฐานของการศึกษาค้นคว้า

1. จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา ด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน มีจำนวนข้อต่างกัน

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา ด้านตัวเลข หลังจากคัดเลือกข้อที่ทำหน้าที่ต่างกันออก ด้วยวิธีการตรวจสอบการทำหน้าที่ต่างกัน มีค่าความเชื่อมั่นต่างกัน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า ผู้วิจัยได้แบ่งออกเป็นหัวข้อ ดังนี้

1. เอกสารที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ
2. เอกสารที่เกี่ยวกับความหมายและทฤษฎีที่เกี่ยวข้องกับความถนัด
3. งานวิจัยที่เกี่ยวข้อง

1. เอกสารที่เกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ

1.1 ความเป็นมาและความหมายของการทำหน้าที่ต่างกันของข้อสอบ

การที่นำข้อสอบหรือแบบทดสอบไปใช้กับผู้สอบที่มีความสามารถหลัก (Primary Ability) ที่ต้องการวัดหรือคุณลักษณะแฝงเป้าหมายของการวัดเท่ากัน แต่มีคุณลักษณะแฝงอื่นต่างกันแล้วทำให้ผู้ที่มาจากต่างกลุ่มกันมีโอกาสในการตอบข้อสอบได้ถูกต้องไม่เท่ากันซึ่งความแตกต่างที่เกิดขึ้นอาจเกิดจากข้อคำถาม ประสิทธิภาพ หรือพื้นฐานเดิมที่แตกต่างกันของผู้สอบ นั้นเดิมเรียกว่า “ ความลำเอียงของข้อสอบ (Item Bias) ” หรือที่ปัจจุบันเรียกว่า “ การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning) ” เป็นคำที่ใช้ในการศึกษาหรือทดสอบความยุติธรรมของข้อสอบ ดังที่ คามิลลีและเชฟพิด (Camilli and Shepard.1994) ได้เสนอแนวความคิดว่า การทำหน้าที่ต่างกันของข้อสอบเน้นที่วิธีการทางสถิติที่นำมาตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันจากกลุ่มผู้สอบต่างกลุ่มกัน ส่วนข้อสอบที่มีความลำเอียงนั้นเป็นกลุ่มข้อสอบย่อยที่นำมาศึกษาโดยไม่คำนึงถึงวิธีการสถิติ ซึ่งมีผู้ให้ความหมายของ “ การทำหน้าที่ต่างกันของข้อสอบ ” ดังนี้

เคเดอร์แมนและเกียร์ดีย์ (Kelderman and Georerdy.1990:370) ได้ให้ความหมายว่า การทำหน้าที่ต่างกันของข้อสอบ เป็นคะแนนข้อสอบที่ได้จากกลุ่มผู้สอบที่มีความสามารถเท่ากันแต่มาจากต่างกลุ่มกันมีความแตกต่างกันอย่างมีระบบ

คามิลลีและเชฟพิด (Camilli and Shepard.1994:16) ได้ให้ความหมายว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นการตรวจสอบความเป็นพหุมิติของข้อสอบ โดยจะแสดงการทำหน้าที่ต่างกันระหว่างกลุ่มผู้สอบตั้งแต่สองกลุ่มขึ้นไปที่มีความสามารถหลัก (Primary Abilities) เท่ากัน แต่มีความสามารถรอง (Secondary Abilities) แตกต่างกัน

โปเทนซ่าและโดแรนส์ (Potenza and Dorans.1995:23) ได้ให้ความหมายว่า การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ผลการตอบข้อสอบระหว่างกลุ่มผู้สอบสองกลุ่มที่นำมาเปรียบเทียบมีความแตกต่างกัน การเปรียบเทียบกลุ่มผู้สอบเป็นสิ่งสำคัญที่จะอธิบายถึงความแตกต่างระหว่างการทำหน้าที่ของข้อสอบกับคุณลักษณะแฝงของกลุ่มผู้สอบ

นารายานาน และสวามินาธาน (Narayanan and Swaminathan.1996:) ได้ให้ความหมายว่า การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ผู้สอบมีความสามารถระดับเดียวกันแต่มาจากกลุ่มย่อยต่างกัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน

ดังนั้น จึงสรุปได้ว่า การทำหน้าที่ต่างกันของข้อสอบ เกิดขึ้นเมื่อข้อสอบหรือแบบทดสอบวัดคุณลักษณะแฝงอื่นนอกเหนือจากคุณลักษณะแฝงที่ต้องการวัด ทำให้ผู้สอบแต่ละกลุ่มที่นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกันทั้ง ๆ ที่มีความสามารถที่ต้องการวัดเท่ากัน

1.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มี 2 รูปแบบ (Camilli and Shepard.1994:59-60) คือ

1. การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform DIF) หมายถึง การที่ข้อสอบทำให้ผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบได้ถูกต้องมากกว่าอีกกลุ่มหนึ่งสม่ำเสมอในทุกระดับความสามารถ โดยเมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบทั้งสองกลุ่มจะพบว่าไม่มีปฏิสัมพันธ์ระหว่างโค้งคุณลักษณะข้อสอบในทุกระดับความสามารถ

2. การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป (Nonuniform DIF) หมายถึง การที่ข้อสอบทำให้โอกาสในการตอบข้อสอบได้ถูกต้องของผู้สอบระหว่างกลุ่มไม่สม่ำเสมอในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของทั้งสองกลุ่มจะพบว่าไม่มีปฏิสัมพันธ์ระหว่างโค้งคุณลักษณะข้อสอบของทั้งสองกลุ่มจะพบว่าไม่มีปฏิสัมพันธ์ระหว่างโค้งคุณลักษณะข้อสอบ เช่น ที่ความสามารถระดับหนึ่ง กลุ่ม A มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่ม B แต่ที่ความสามารถอีกระดับหนึ่ง กลุ่ม B กลับมีโอกาสในการตอบข้อสอบได้ถูกมากกว่ากลุ่ม A

1.3 วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีด้วยกันหลายวิธี โดยมีการศึกษา และคิดค้นวิธีการต่าง ๆ เพื่อให้การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้อย่างมีประสิทธิภาพมากขึ้น ซึ่งสามารถแบ่งตามประเภทการวิเคราะห์ ดังนี้ (Potenza and Dorans.1995:28-35)

กลุ่มวิธีทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นวิธีการวิเคราะห์ความแตกต่างของฟังก์ชันการตอบสนองข้อสอบโดยเปรียบเทียบโค้งคุณลักษณะข้อสอบของกลุ่มผู้สอบตามระดับความสามารถผู้สอบ ถ้าโค้งคุณลักษณะข้อสอบของกลุ่มผู้สอบ 2 กลุ่มมีรูปร่างเหมือนกัน แสดงว่าข้อสอบข้อนั้นทำหน้าที่ไม่ต่างกัน แต่ถ้าโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่มมีรูปร่างแตกต่างกัน แสดงว่าข้อสอบข้อนั้นทำหน้าที่ต่างกัน กระบวนการวิเคราะห์ใช้สถิติพาราเมตริก มีดัชนีบอกระดับการทำหน้าที่ต่างกันและทดสอบนัยสำคัญทางสถิติ ซึ่งได้รับการพัฒนาและแตกย่อยออกไปหลายวิธี เช่น วิธี General IRTLR วิธี Loglinear IRTLR วิธี IRT-D² และวิธี Lord's χ^2 จากการศึกษาพบว่าวิธีนี้เป็นวิธีที่มีความถูกต้องสูง สามารถตรวจสอบพบจำนวนข้อสอบที่ทำหน้าที่ต่างกันมากที่สุด (Holland and Thayer.1988) ค่าสถิติของข้อสอบไม่เปลี่ยนแปลงไปตามกลุ่มตัวอย่าง และการใช้การประมาณค่าความสามารถที่แท้จริงของผู้สอบแทนคะแนนที่สังเกตได้ แต่มีข้อจำกัด คือ ต้องใช้กลุ่มตัวอย่างขนาดใหญ่ ข้อมูลต้องเป็นไปตามข้อตกลงเบื้องต้น มีการคำนวณที่ซับซ้อน แปลผลยาก ค่าใช้จ่ายสูง

วิธีถดถอยโลจิสติก (Logistic Regression:LR) เป็นวิธีการที่สามารถตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบได้ถูกต้องใกล้เคียงกับวิธีทฤษฎีการตอบสนองข้อสอบ มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบแบบเอกรูป (Uniform DIF) ได้เท่าวิธีแมนเทิล-แฮนดเชล และวิธีซิบเทสต์ (Swaminatan and Roger .1990) แล้วยังสามารถตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบอเนกรูป (Nonuniform DIF) ได้ดีเช่นเดียวกัน

วิธีซิบเทสต์ (SIBTEST) เป็นวิธีที่มีพื้นฐานอยู่บนทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ใช้คะแนนรวมจากแบบทดสอบเป็นตัวแทนความสามารถโดยมีข้อตกลงเบื้องต้นว่ามีมิติการวัด 2 มิติ คะแนนรวมจากแบบทดสอบจึงมี 2 ส่วน คือ คะแนนแบบทดสอบที่มีความเที่ยงตรง (Valid Subtest) ซึ่งวัดคุณลักษณะแฝง

เป้าหมาย และคะแนนรวมจากแบบทดสอบที่ศึกษา (Studied Subtest) ซึ่งวัดคุณลักษณะแฝงแทรกซ้อน มีดัชนีบอกระดับการทำหน้าที่ต่างกันของข้อสอบ และการทดสอบนัยสำคัญทางสถิติ วิธีซิปเทสต์นี้มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบเอกรูปได้ใกล้เคียงกับ

วิธีแมนเทล-แฮนดเซล (Shealy and Stout.1993)

ในการวิจัยครั้งนี้ผู้วิจัยได้เลือกที่จะใช้วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดย วิธีแมนเทล-แฮนดเซล และวิธีวิเคราะห์ความแปรปรวน ซึ่งมีรายละเอียด ดังนี้

1.3.1 วิธีแมนเทล-แฮนดเซล (Mantel-Haenszel)

วิธีแมนเทล-แฮนดเซล เป็นวิธีที่ Mantel และ Haenszel ได้เสนอขึ้นใช้ในปี 1959 โดยพัฒนา มาจากวิธีไคสแควร์ซึ่งเป็นวิธีที่อาศัยทฤษฎีการวัดแบบดั้งเดิมเป็นพื้นฐานโดยมี Holland and Thayer ได้เริ่ม นำมาใช้เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ซึ่งต่อมาถูกนำมาใช้อย่างกว้างขวางเนื่องจากเป็นวิธี ที่ใช้ง่าย สะดวก และประหยัด (Holland and Thayer.1988) หลักการตรวจสอบการทำหน้าที่ต่างกันของ วิธีแมนเทล-แฮนดเซล เป็นการเปรียบเทียบผลการตอบข้อสอบของผู้สอบ 2 กลุ่ม ได้แก่ กลุ่มอ้างอิง (Reference group:R) ซึ่งเป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบ และกลุ่มเปรียบเทียบ (Focal group:F) เป็นกลุ่มที่เสียประโยชน์จากการตอบข้อสอบ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แต่ละครั้งจะเรียกข้อสอบที่ถูกตรวจว่า ข้อสอบที่ศึกษา (studied item) โดยจะเปรียบเทียบในทุกๆ ระดับ คะแนนรวมที่ได้จากผลการตอบข้อสอบของผู้สอบ ข้อสอบใดที่ผู้สอบในกลุ่มที่มีความสามารถเท่ากันทำได้ถูกต้องเท่ากัน แสดงว่า เป็นข้อสอบที่ทำหน้าที่ไม่ต่างกัน (no DIF) ระหว่างกลุ่มผู้สอบ การตรวจสอบด้วยวิธีนี้ทำได้โดยแยกวิเคราะห์ข้อสอบเป็นรายข้อ ซึ่งในแต่ละข้อจะต้องสร้างตารางการถ่วงแบบ 2x2 (กลุ่มผู้สอบ 2 กลุ่ม x ผลการตอบ 2 แบบ) แสดงความถี่ของผู้สอบที่ตอบถูก/ผิด ในกลุ่มอ้างอิง (R) และกลุ่มเปรียบเทียบ (F) ตามช่วงคะแนนของผู้สอบสำหรับช่วงคะแนน j ถ้ามีการแบ่งช่วงคะแนนการสอบเป็น k ช่วง จะได้ ตารางการถ่วงแบบ 2x2 จำนวน k ตาราง

ตาราง 1 แสดงความถี่ของผู้สอบที่ตอบข้อสอบถูกและผิดในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบสำหรับช่วงคะแนน j

กลุ่มผู้สอบ	คะแนนที่ได้จากข้อสอบที่ต้องการวิเคราะห์ DIF		
	ตอบถูก(1)	ตอบผิด(0)	รวม
กลุ่มอ้างอิง R	A_j	B_j	n_{Rj}
กลุ่มเปรียบเทียบ F	C_j	D_j	n_{Fj}
รวม	m_{1j}	m_{0j}	T_j

เมื่อ	A_j	เป็นความถี่ที่สังเกตได้ในการตอบถูกในช่วงคะแนน j ของกลุ่มอ้างอิง
	B_j	เป็นความถี่ที่สังเกตได้ในการตอบผิดในช่วงคะแนน j ของกลุ่มอ้างอิง
	C_j	เป็นความถี่ที่สังเกตได้ในการตอบถูกในช่วงคะแนน j ของกลุ่มเปรียบเทียบ
	D_j	เป็นความถี่ที่สังเกตได้ในการตอบผิดในช่วงคะแนน j ของกลุ่มเปรียบเทียบ
	m_{1j}	จำนวนผู้สอบที่ตอบถูกทั้งหมดในช่วงคะแนน j
	m_{0j}	จำนวนผู้สอบที่ตอบผิดทั้งหมดในช่วงคะแนน j

T_j จำนวนผู้สอบทั้งหมด
โดยแสดงเป็นสัดส่วนการตอบข้อสอบของผู้สอบ 2 กลุ่ม ได้ดังตาราง 2

ตาราง 2 แสดงสัดส่วนการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่คะแนน j

ผลการตอบของกลุ่ม	คะแนน		
	1	0	รวม
กลุ่มอ้างอิง R	P_{Rj}	q_{Rj}	1
กลุ่มเปรียบเทียบ F	P_{Fj}	q_{Fj}	1

เมื่อ P_{Rj} คือ สัดส่วนของกลุ่มอ้างอิงที่ตอบข้อสอบถูก (1) ณ ระดับชั้นคะแนน j
 P_{Fj} คือ สัดส่วนของกลุ่มเปรียบเทียบที่ตอบข้อสอบถูก (1) ณ ระดับชั้นคะแนน j
 q_{Rj} คือ สัดส่วนของผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบผิด (0) ณ ระดับชั้นคะแนน j ($q_{Rj}=1-P_{Rj}$)
 q_{Fj} คือ สัดส่วนของผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบผิด (0) ณ ระดับชั้นคะแนน j ($q_{Fj}=1-P_{Fj}$)

หลักการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนด์เชลเป็นการนำข้อมูลจากตารางไขว้ k ตารางมาทำการวิเคราะห์ตามขั้นตอน ดังนี้

1. คำนวณค่าความน่าจะเป็นรูปในรูปของสัดส่วนของการตอบข้อสอบถูกและผิดระหว่างกลุ่ม ในทุกช่วงคะแนน จากสูตร

$$\alpha_{MH} = \frac{\sum_{j=1}^s A_j D_j / T_j}{\sum_{j=1}^s B_j C_j / T_j}$$

เมื่อ A_j แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบถูกในช่วงคะแนน j
 B_j แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบผิดในช่วงคะแนน j
 C_j แทน ผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบถูกในช่วงคะแนน j
 D_j แทน ผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบผิดในช่วงคะแนน j
 T_j แทน ผู้สอบทั้งสองกลุ่มที่ได้รับคะแนนรวมจากการสอบในช่วงคะแนน j
 α_{MH} แทน สัดส่วนของการตอบข้อสอบถูกและผิดระหว่างกลุ่มในแต่ละข้อใน

ทุกช่วงคะแนน j

โดยพิจารณาค่า α_{MH} มีค่าระหว่าง 0 ถึง ∞ เกณฑ์ในการพิจารณาข้อสอบทำหน้าที่ต่างกัน

-ถ้าค่า α_{MH} ที่คำนวณได้มีค่าเท่ากับ 1 แสดงว่าข้อสอบทำหน้าที่ไม่แตกต่างกัน

-ถ้าค่า $\alpha_{MH} > 1$ แสดงว่าข้อสอบเข้าข้างกลุ่มอ้างอิง

2. ทดสอบนัยสำคัญของค่าสถิติไคสแควร์ เพื่อทดสอบค่า α_{MH} ที่คำนวณได้ว่าแตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 หรือไม่ ที่ระดับชั้นความเป็นอิสระเท่ากับ 1 ตามสูตร ดังนี้

$$\chi^2 = \frac{\left\{ \sum A_{ij} - \sum E(A_{ij}) - 0.5 \right\}^2}{\sum \text{Var}(A_{ij})}$$

$$\text{เมื่อ } E(A_{ij}) = (n_{.j})(m_{i.}) / T_j$$

$$\text{Var}(A_{ij}) = \frac{n_{.j} n_{i.} m_{i.} m_{.j}}{T_j^2 (T_j - 1)}$$

ดังนั้นการตรวจสอบด้วยวิธีแมนเทิล-แฮนดิล ด้วยสมมติฐานหลักที่แสดงว่าข้อสอบข้อนั้นทำหน้าที่ไม่ต่างกันดังนี้ คือ

$$H_0 : P_{Rj} = P_{Fj} \quad \text{สำหรับทุกชั้นคะแนน } j$$

สมมติฐานหลักนี้เป็นสมมติฐานของความเป็นอิสระอย่างมีเงื่อนไขของสมาชิกกลุ่มและคะแนนที่ได้จากการตอบข้อสอบที่ต้องการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ และภายใต้สมมติฐานหลักจะได้ค่าคาดหวังของการตอบ ดังนี้

$$E(A_j) = (n_{Rj} m_{1.}) / T_j$$

$$E(B_j) = (n_{Rj} m_{2.}) / T_j$$

$$E(C_j) = (n_{Rj} m_{3.}) / T_j$$

$$E(D_j) = (n_{Rj} m_{4.}) / T_j$$

สมมติฐานอื่นของวิธีแมนเทิล-แฮนดิล คือ

$$H_1 : \frac{p_{Rj}}{q_{Rj}} = \alpha \frac{p_{Fj}}{q_{Fj}} \quad j=1, \dots, k \quad \text{เมื่อ } \alpha \text{ ไม่เท่ากับ } 1$$

แต่เมื่อ α เท่ากับ 1 ซึ่งสอดคล้องกับสมมติฐานศูนย์ จะได้ว่า

$$H_0 : \frac{p_{Rj}}{q_{Rj}} = \frac{p_{Fj}}{q_{Fj}} \quad j = 1, \dots, k$$

และค่าประมาณของ

$$\alpha_{MH} = \frac{p_{Rj}}{q_{Rj}} / \frac{p_{Fj}}{q_{Fj}} = \frac{p_{Rj} q_{Fj}}{p_{Fj} q_{Rj}} \quad \text{สำหรับทุก } j = 1, \dots, k$$

3. Holland และ Thayer ได้เสนอแนะให้แปลงค่า α_{MH} ให้เป็นค่าเดลต้า (Δ_{MH}) หรือ MH_{DIF} ดังนี้

$$MH_{DIF} = -2.35 \ln(\alpha_{MH})$$

สำหรับเกณฑ์ในการตัดสินข้อสอบที่ทำหน้าที่ต่างกัน พิจารณาจากข้อสอบที่มีค่า α_{MH} แตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติ หรือค่า MH_{DIF} แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติและมีเกณฑ์ในการพิจารณาค่าของ MH_{DIF} ดังนี้

1. ค่า MH_{DIF} เท่ากับ 0 หรือไม่แตกต่างจาก 0 อย่างมีนัยสำคัญแสดงว่า ข้อสอบข้อนั้นทำหน้าที่ไม่แตกต่างกันระหว่างกลุ่ม (No DIF)

2. ค่า MH_{DIF} แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ และมีค่าเป็นบวก แสดงว่าข้อสอบข้อนั้นทำหน้าที่ต่างกัน โดยจะลำเอียงเข้าหากกลุ่มเปรียบเทียบ

3. ค่า MH_{DIF} แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ และมีค่าเป็นลบ แสดงว่าข้อสอบนั้น ทำหน้าที่ต่างกัน โดยจะลำเอียงเข้าหากกลุ่มอ้างอิง

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนด์เชลล์ เป็นการเปรียบเทียบผลการตอบข้อสอบของผู้สอบสองกลุ่ม หลังจากการจับคู่ผู้สอบตามความรู้ หรือความสามารถของผู้สอบ การจับคู่ผู้สอบของผู้สอบสองกลุ่มเป็นเงื่อนไขที่สำคัญเพราะเป็นเกณฑ์ที่ใช้แทนความสามารถที่แท้จริงของผู้สอบ สองกลุ่ม ในทางปฏิบัติมักใช้คะแนนรวมของแบบทดสอบ (Total Test Score) เป็นเกณฑ์ในการจับคู่ เพราะเกี่ยวข้องกับความรู้ หรือ ความสามารถที่วัดได้จากแบบทดสอบนั้น และจากผู้สอบทุกคนภายใต้สถานการณ์เดียวกัน แต่การใช้คะแนนรวมของแบบทดสอบเป็นเกณฑ์ในการจับคู่ผู้สอบมีจุดอ่อน คือ มีการรวมเอาคะแนนจากข้อสอบที่ทำหน้าที่ต่างกัน มาเป็นเกณฑ์ในการจับคู่ผู้สอบด้วย จุดอ่อนนี้สามารถแก้ไขได้โดยใช้การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบ 2 ขั้นตอน ซึ่ง Holland และ Thayer (1988) ได้เสนอขั้นตอนการวิเคราะห์ ดังนี้

ขั้นตอนที่ 1 ใช้คะแนนรวมของแบบทดสอบทั้งฉบับเป็นเกณฑ์ในการจับคู่ผู้สอบ แล้ววิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ

ขั้นตอนที่ 2 นำคะแนนของข้อสอบที่ทำหน้าที่ต่างกันที่ตรวจพบในขั้นตอนที่ 1 ออกจากเกณฑ์การจับคู่ผู้สอบแล้วใช้คะแนนรวมของข้อสอบที่ทำหน้าที่ไม่แตกต่างกัน หรือข้อสอบที่เหลือเป็นเกณฑ์การจับคู่ผู้สอบในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบของแบบทดสอบทั้งฉบับในขั้นตอนที่ 2

1.3.2. วิเคราะห์ความแปรปรวน (Analysis of Variance)

การวิเคราะห์ความแปรปรวน เป็นวิธีหนึ่งที่สามารถการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบได้ โดยพิจารณาว่าข้อสอบที่ทำหน้าที่ต่างกันโดยวัดความสามารถเดียวกันจากผู้สอบต่างกลุ่มกัน เช่น ข้อสอบที่ง่ายสำหรับกลุ่มหนึ่ง แต่ยากสำหรับอีกกลุ่มหนึ่ง แสดงว่าข้อสอบข้อนั้นมีการทำหน้าที่ต่างกันของ

ข้อสอบโดยทำการวิเคราะห์ความแปรปรวนก่อนแล้วทำการทดสอบนัยสำคัญทางสถิติของผลปฏิสัมพันธ์ (Interaction Effect) ทั้งหมดระหว่างข้อสอบกับกลุ่มผู้สอบ ซึ่งมีรูปแบบการวิเคราะห์ ดังนี้ (Laksana.1979:12)

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}$$

เมื่อ	μ	แทน ค่าเฉลี่ยทั้งหมด
	α_i	แทน อิทธิพลจากข้อสอบข้อที่ i (i=1,2,3,...,i)
	β_j	แทน อิทธิพลจากกลุ่มผู้สอบ กลุ่มที่ j (j=1,2,3,...,j)
	$(\alpha\beta)_{ij}$	แทน ปฏิสัมพันธ์ระหว่างข้อสอบกับกลุ่มผู้สอบ
	ε_{ij}	แทน ความคลาดเคลื่อนแบบสุ่มภายในเซลล์ i,j

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยวิธีการวิเคราะห์ความแปรปรวนนี้มีข้อตกลงว่า

1. แบบทดสอบจะต้องมีลักษณะเป็นเอกพันธ์ คือต้องวัดความสามารถ หรือคุณลักษณะเพียงลักษณะเดียว

2. ความแปรปรวนของข้อสอบแต่ละข้อต้องเท่ากัน

3. ความแปรปรวนร่วมของข้อสอบแต่ละคู่ต้องเท่ากัน

ขั้นตอนการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีวิเคราะห์ความแปรปรวน

1. จัดข้อสอบ (Item) และกลุ่มผู้สอบ (Subgroups) ลงในตัวแปรผลการตอบข้อสอบ และกลุ่มผู้สอบ

2. คำนวณสัดส่วนในการตอบข้อสอบแต่ละข้อถูกระหว่างกลุ่มทั้งสอง

3. ทดสอบ Overall Interaction Effect ระหว่างข้อ (Item) กับกลุ่มผู้สอบ (Subgroups) หากพบว่า Interaction Effect มีนัยสำคัญทางสถิติ (Significant) แสดงว่าแบบทดสอบฉบับนั้นมีการทำหน้าที่ต่างกันของข้อสอบ

4. หากการทดสอบในข้อ 3 มีนัยสำคัญทางสถิติ ให้ทำการทดสอบต่อเป็นรายข้อโดยใช้การทดสอบ t-test ถ้าหากค่า t ที่คำนวณได้มีค่าเท่ากับหรือมากกว่า ค่าวิกฤตของ t ในตาราง (t = 1.96) ถือว่าข้อสอบข้อนั้นมีการทำหน้าที่ต่างกันของข้อสอบ

2. เอกสารที่เกี่ยวกับความหมายและทฤษฎีที่เกี่ยวข้องกับความถนัด

2.1 ความหมายของความถนัด

เนื่องจากมนุษย์มีความแตกต่างกันทั้งร่างกาย จิตใจ และสติปัญญา มนุษย์แต่ละคนจึงมีความถนัดที่แตกต่างกันไป นักจิตวิทยาหลายท่าน และนักวัดผล ได้กล่าวถึงความถนัดในความหมายต่างๆ ไว้ดังนี้

วอร์เรน (Warren.1934:18) ให้ความหมายว่า ความถนัดเป็นสภาวะหรือคุณลักษณะที่จะแสดงความสามารถของแต่ละบุคคลในการเรียนรู้ ซึ่งได้มาจากการฝึกฝนทักษะหรือการตอบสนองต่อเรื่องใดเรื่องหนึ่งโดยเฉพาะ

บิงแฮม (Bingham.1937:18) ให้ความหมายว่า ความถนัดเป็นสภาวะอันแสดงความเหมาะสมของบุคคลที่สำคัญคือการเพิ่มความชำนาญให้กับตนเองหรือเป็นศักยภาพของบุคคลและรวมถึงความพร้อมที่จะสนใจในความสามารถนั้น

ครอนบาค (Cronbach.1970: 38) ให้ความหมายว่า เป็นความสามารถทางสมองที่ร่วมกันทำงานเพื่อเพิ่มพูนความสำเร็จในกิจกรรมทางปัญญา

ชวาล แพรัตกุล (2517:50) ให้ความหมายว่า เป็นสมรรถวิสัยและทิศทางแห่งความงอกงามของสมอง หรือกล่าวให้ง่ายขึ้นก็หมายถึงขีดความสามารถสูงสุดของบุคคลที่เขาจะมีได้ต่อการเรียนรู้ การฝึกฝน วิทยาการต่างๆ และทักษะทั้งปวง ถ้าเขาได้รับการฝึกฝนและประสบการณ์ที่เหมาะสม

บุญส่ง นิลแก้ว.(2517:177) ให้ความหมายว่า เป็นสมรรถภาพหรือศักยภาพที่มีอยู่ในตัวบุคคล อันจะก่อให้เกิดความสำเร็จในการทำกิจกรรมใดๆ

ชาญวิทย์ เทียมบุญประเสริฐ (2528:14) ให้ความหมายว่า ความถนัดเป็นสมรรถภาพพื้นฐานทางสมองที่มีอยู่ในตัวบุคคลมีลักษณะเป็นกลุ่มของสมรรถภาพ ซึ่งเกิดจากการสั่งสม ประสบการณ์ที่มีผลต่อการเรียนรู้ และความสำเร็จในอนาคต

ล้วน สายยศ และอังคณา สายยศ (2541:15-16) ให้ความหมายว่า เป็นความสามารถทางสมองที่บุคคลได้รับจากการฝึกฝน และมีการสั่งสมไว้มากจนเกิดเป็นทักษะเด่นชัดด้านใดด้านหนึ่ง

ดังนั้นจึงกล่าวได้ว่า ความถนัด หมายถึง ความสามารถของแต่ละบุคคลที่สามารถเรียนรู้และฝึกฝนให้เกิดความชำนาญ โดยแยกออกเป็นความสามารถในด้านต่างๆ อันจะนำมาเพื่อความสำเร็จในงาน

2.2 ทฤษฎีเกี่ยวกับความถนัด

ความถนัดเป็นกระบวนการทางสมองของมนุษย์ที่ซับซ้อน เป็นองค์ประกอบที่สำคัญยิ่งในการที่บุคคลจะประสบความสำเร็จในกิจกรรมต่างๆ ดังนั้นนักจิตวิทยา และนักการศึกษาหลายท่านได้พยายามศึกษาค้นคว้าเกี่ยวกับความถนัด จนกระทั่งนำมาสร้างเป็นทฤษฎีต่างๆ เพื่อใช้ในการอธิบายความสามารถทางสมองของบุคคลได้ต่างๆ กัน ดังนี้ (ล้วน สายยศ และอังคณา สายยศ.2541:43-56)

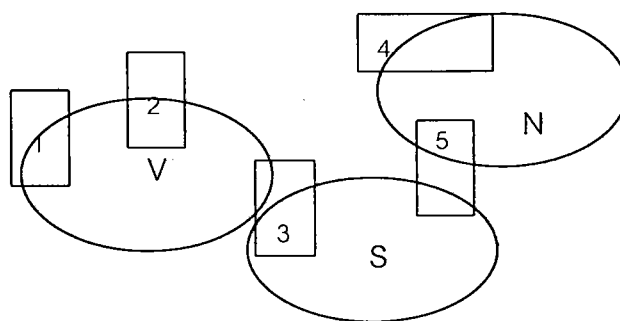
2.2.1. ทฤษฎีองค์ประกอบเดียว (Uni-factor Theory)

เป็นทฤษฎีที่คิดขึ้นโดยบิเนต์และไซมอน (Binet and Simon) ในปีค.ศ.1905 โดยเสนอโครงสร้างเป็นอันหนึ่งอันเดียวกันไม่แบ่งแยกออกเป็นส่วนย่อยคล้ายกับเป็นความสามารถทั่วไป (General Ability) และวัดออกมาเป็นคะแนนเดียว

2.2.2. ทฤษฎีสององค์ประกอบ (Bi-factor Thoery)

เป็นทฤษฎีที่คิดขึ้นโดย ชาร์ลส สเปียร์แมน (Charles Spearman) ในปีค.ศ.1927 โดยเป็นทฤษฎีที่เกิดจากการวิเคราะห์คุณลักษณะโดยกระบวนการทางสถิติ พบว่ากิจกรรมทางสมองทั้งหลายเมื่อวิเคราะห์ดูแล้วมีองค์ประกอบร่วมอันหนึ่ง เรียกชื่อองค์ประกอบนี้ว่า องค์ประกอบทั่วไป (General Factor : G-Factor) เนื่องจากหาสหสัมพันธ์เกี่ยวพันกันในแต่ละองค์ประกอบ (Intercorrelation) มีค่าสูง แต่สูงอย่างไม่สมบูรณ์แบบ จึงให้ชื่อองค์ประกอบอื่นย่อยๆ นี้ว่า องค์ประกอบเฉพาะ (Specific Factor : S-Factor) แต่ละองค์ประกอบนี้จะมีกิจกรรมเฉพาะตัวชนิดหนึ่งและจะเห็นว่าทฤษฎีนี้มองความสำคัญที่องค์ประกอบต่างๆ ไปเป็นหลัก ไม่แตกต่างจากทฤษฎีองค์ประกอบเดียวของบิเนต์ ส่วนที่แตกต่างก็คือนอกจากองค์ประกอบร่วมแล้วยังมีองค์ประกอบย่อยเพิ่มขึ้นอีกซึ่งเป็นแนวคิดใหม่ที่ไม่เคยปรากฏมาก่อนนี้

2.2.3. ทฤษฎีหลายองค์ประกอบ (Multiple-factor Theory) เป็นทฤษฎีที่ยอมรับกันอย่างกว้างขวางของนักจิตวิทยาชาวอเมริกัน ผู้นำในการสร้างทฤษฎีนี้ คือ เทอร์สโตน (L.L.Thurstone.1933) ได้ทำการวิจัยโครงสร้างทางสมองอย่างกว้างขวางและใช้หลักการวิเคราะห์สมัยใหม่ที่ เรียกว่าการวิเคราะห์องค์ประกอบ (Factor Analysis) ทำให้สามารถแยกความสามารถทางสมองออกเป็นส่วนย่อยๆ ได้หลายองค์ประกอบโดยเชื่อว่าความสามารถทางสมองไม่ได้ประกอบด้วยความสามารถรวมเป็นแกนกลาง แต่ประกอบด้วยองค์ประกอบเป็นกลุ่มๆ หลายๆ กลุ่ม โดยแต่ละกลุ่มจะมีหน้าที่เป็นอย่างไรๆ โดยเฉพาะ หรืออาจทำงานร่วมกันบ้างก็ได้ องค์ประกอบย่อยๆ นี้ เทอร์สโตนให้ชื่อว่า ความสามารถปฐมภูมิของสมอง (Primary Mental Abilities) แยกองค์ประกอบย่อยโดยยึดน้ำหนักขององค์ประกอบเด่นๆ (Loading Factor) เป็นสำคัญ โดยกลุ่มของความสามารถ หรือองค์ประกอบก็ยังทำหน้าที่เกี่ยวพันกันบ้างเหมือนกัน ดังเช่น องค์ประกอบด้านภาษา (Verbal Factor) เป็นน้ำหนักองค์ประกอบมากที่สุด คือ ความสามารถทางศัพท์ น้ำหนักลดลงมาอีก คือ คณิตศาสตร์เหตุผล ซึ่งสามารถแสดงความสัมพันธ์ภายในแบบทดสอบ 5 ชุด ที่ขึ้นอยู่กับ 3 องค์ประกอบ คือ V : Verbal, N : Number และ S : Spatial ตามทฤษฎีหลายองค์ประกอบดังนี้ (Anastasi. 1988:383)



ภาพประกอบ 1 โครงสร้างความสามารถทางสมองของทฤษฎีหลายองค์ประกอบ

จากภาพประกอบ 1 แสดงให้เห็นว่าสหสัมพันธ์ของแบบทดสอบ 1 , 2 และ 3 ที่มีต่อกันนั้นมีองค์ประกอบร่วมทางภาษา (Verbal Factor : V) ในทำนองเดียวกันสหสัมพันธ์ระหว่างแบบทดสอบ 3 และ 5 เป็นผลจากองค์ประกอบมิติสัมพันธ์ (Spatial Factor : S) (หรือ) และความสัมพันธ์ระหว่างแบบทดสอบ 4 และ 5 เป็นผลจากองค์ประกอบด้านตัวเลข (Numeric Factor : N) (หรือ) ที่น่าสนใจคือ แบบทดสอบ 3 และ 5 มีองค์ประกอบซ้อนขึ้นมา นั่นคือ V และ S มีอยู่ในแบบทดสอบ 3, N และ S มีอยู่ในแบบทดสอบ 5

เทอร์สโตนวิเคราะห์องค์ประกอบความสามารถของมนุษย์ได้หลายองค์ประกอบ แต่ที่เห็นได้ชัด และสำคัญมี 7 องค์ประกอบ คือ

1. องค์ประกอบด้านภาษา (Verbal Factor : V) เป็นความสามารถด้านการเข้าใจในภาษา และการสื่อสารทั่วไป ผู้มีองค์ประกอบด้านนี้สูงจะมีความสามารถอ่านเอาเรื่อง อ่านแบบเข้าใจความหมาย รู้ความสัมพันธ์ของคำ รู้ความหมายของคำศัพท์ได้เป็นอย่างดี

2. องค์ประกอบด้านความคล่องแคล่วในการใช้ถ้อยคำ (Word Fluency Factor : W) เป็นความสามารถที่จะใช้คำได้มากในเวลาจำกัด เช่น ให้หาคำขึ้นต้นด้วย " ก " ให้มากที่สุดในเวลาจำกัด เป็นต้น ความสามารถด้านนี้จะส่งผลให้มีความสามารถในการเจรจา และการประพันธ์ทั้งร้อยแก้ว

และร้อยกรอง ตอบโต้ทันทีทันใด หรือ เรียกว่ามีปฏิภาณไหวพริบในการเจรจา ซึ่งต่างจากองค์ประกอบด้านภาษาที่มองเป็นความสามารถด้านในภาษาทางความคิด ความเข้าใจทางภาษา ส่วนองค์ประกอบด้านความคล่องแคล่วในการใช้ถ้อยคำ มองเป็นผลในด้านการเจรจา เปรียบเทียบได้กับคนที่เขียนเก่ง (V) แต่พูดบรรยาย (W) ให้คนอื่นฟังไม่รู้เรื่อง

3. องค์ประกอบด้านจำนวน (Number Factor : N) องค์ประกอบนี้ส่งผลให้มีความเข้าใจในวิชา คณิตศาสตร์ต่างๆ ได้ดี มีความสามารถมองเห็นความสัมพันธ์และความหมายของจำนวน และมีความแม่นยำคล่องแคล่วในการบวก ลบ คูณ หาร ในวิชา เลขคณิตได้อย่างดีด้วย

4. องค์ประกอบด้านมิติสัมพันธ์ (Space Factor : S) ด้านนี้จะส่งผลให้คนเข้าใจถึงขนาดและมิติต่างๆ อันได้แก่ ความสั้น ยาว ใกล้เคียง และพื้นที่ หรือทรวดทรงที่มีขนาดและปริมาตรแตกต่างกัน สามารถสร้างจินตนาการให้เห็นส่วนย่อยและส่วนผสมของวัตถุต่างๆ เมื่อนำเข้ามาซ้อนทับกัน สามารถรู้ความสัมพันธ์ของรูปทรงเรขาคณิตเมื่อเปลี่ยนแปลงที่อยู่ได้

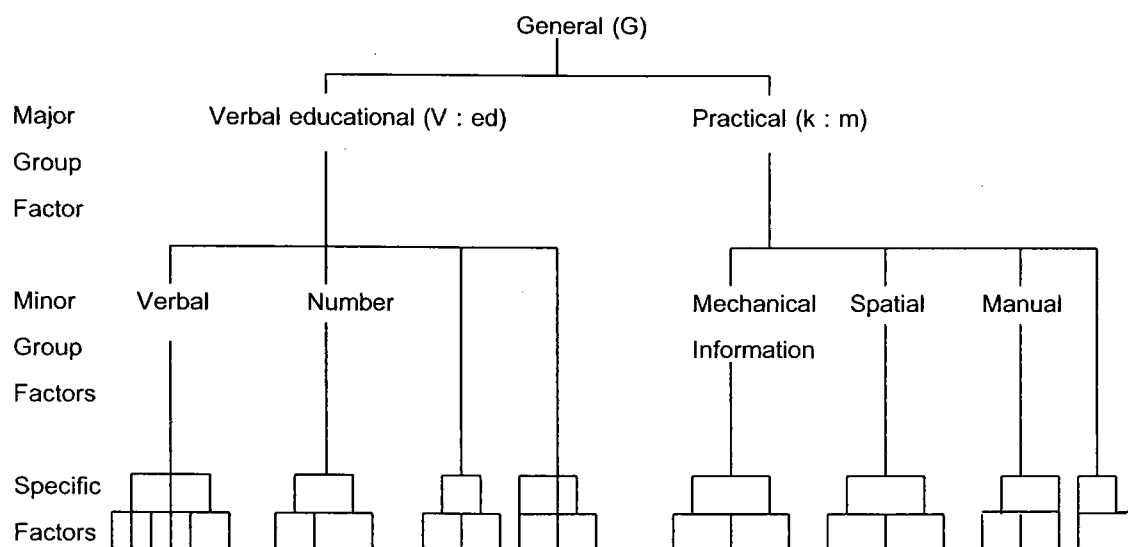
5. องค์ประกอบด้านความจำ (Memory Factor : M) เป็นความสามารถด้านความทรงจำเรื่องราว และมีสติระลึกจูนสามารถถ่ายทอดได้ เป็นการจำแบบไม่รู้ความหมาย หรือจำโดยอาศัยสิ่งสัมพันธ์ของรูปทรงเรขาคณิต เมื่อเปลี่ยนแปลงที่อยู่

6. องค์ประกอบด้านสังเกตพิจารณา (Perceptual Speed Factor : P) เป็นความสามารถด้านเห็นรายละเอียด ความคล้ายคลึงหรือ ความแตกต่างระหว่างสิ่งของต่างๆ อย่างรวดเร็วและถูกต้อง

7. องค์ประกอบด้านเหตุผล (Reasoning Factor : R) เป็นองค์ประกอบนี้ถึงความสามารถด้านวิจารณ์ญาณในการหาเหตุผล ค้นหาหาความสำคัญ ความสัมพันธ์ และหลักการทั้งหลายที่สร้างกฎหรือ ทฤษฎี มีผู้ทำการศึกษาด้านนี้โดยมองเห็นว่าจะวัดเหตุผลทั่วไปได้ดีด้วยเลขคณิตเหตุผล (Arithmetic Reasoning)

2.2.4. ทฤษฎีไฮราคิคัล (Hierarchical Theories)

ทฤษฎีนี้เป็นทฤษฎีที่เกี่ยวกับความสามารถทางสมอง ที่คิดค้นโดยกลุ่มนักจิตวิทยาชาวอังกฤษ คือ เบิร์ท (Burt) เวอร์นอน (Vernon) และฮัมเฟรย์ (Humphreys) โดยเวอร์นอนได้เสนอโครงสร้างสมรรถภาพทางสมองใหม่อีกรูปแบบหนึ่ง ในปีค.ศ.1960 คือ องค์ประกอบทั่วไป G-Factor เช่นเดียวกับของสเปียร์แมน โดยที่เวอร์นอนได้แบ่ง G-Factor เป็น 2 องค์ประกอบ คือ Verbal education (V : ed) และ Practical mechanical (k : m) องค์ประกอบใหญ่ทั้งสองอันนี้เรียกรวมว่า Major Group Factors และองค์ประกอบใหญ่นี้ยังแบ่งออกเป็นส่วนย่อยๆ ลงไปอีก เป็นองค์ประกอบด้านตัวเลข (Numerical) และ องค์ประกอบด้านภาษา (Verbal) และอื่นๆ อีก ในทำนองเดียวกันองค์ประกอบ Practical-mechanical ยังแบ่งย่อยออกเป็น Mechanical-information , Spatial และ Manual และยังมีอื่นๆ แต่ยังไม่กำหนด กลุ่มองค์ประกอบนี้เรียกว่า Minor Group Factor ระดับที่ต่ำที่สุดขององค์ประกอบในรูปแบบนี้ยังมี องค์ประกอบย่อยๆ ลงไปอีก เรียกว่า องค์ประกอบเฉพาะ (Specific Factor) ถ้าพิจารณาโครงสร้างอันนี้แล้วก็ไม่ต่างกับลักษณะของต้นไม้แม่กิ่ง ก้านใหญ่เล็กลงไปตามลำดับ ลำต้นเปรียบเสมือน (G-Factor) กิ่งก้านเล็กๆ เปรียบเสมือน Specific Factor ดังภาพ (Anastasi.1988:371)



ภาพประกอบ 2 โครงสร้างความสามารถทางสมองของทฤษฎีไฮราคัล

2.2.5. ทฤษฎีโครงสร้างสามมิติของปัญญา (Three Faces of Intellect Model)

กิลฟอร์ด (Guilford) ได้สร้าง Structure of Intellect Model หรือ Three Dimension Model of The Structure of Intellect กิลฟอร์ดได้วิเคราะห์ความสัมพันธ์ของคุณลักษณะโดยจัดระบบของคุณลักษณะให้อยู่ในรูปใหม่เป็นลูกบาศก์รวมกัน 120 ก้อน และนิยามคุณลักษณะเป็น 3 มิติ ดังนี้

มิติที่ 1 ด้านกระบวนการหรือวิธีการของการคิด (Operations) มีส่วนประกอบย่อย 5 ส่วน คือ

1. การรู้จักการเข้าใจ (Cognition) หมายถึง ความสามารถที่เห็นสิ่งเร้าแล้วเกิดการรับรู้เข้าใจในสิ่งนั้นๆ และบอกได้ว่าสิ่งนั้นๆ คืออะไร
2. ความจำ (Memory) หมายถึง ความสามารถในการเก็บสะสมความรู้แล้วสามารถระลึกนึกออกมาได้
3. การคิดนอกเนกนัย (Divergent Production) หมายถึง เป็นความสามารถในการตอบสิ่งเร้าได้หลายแง่หลายมุมแตกต่างกันไป เช่น ให้ออกประโยชน์ของก้อนอิฐมาให้มากที่สุดที่จะบอกได้ ถ้าผู้ใดคิดได้มาก และแปลกที่สุดมีเหตุผล ถือว่าผู้นั้นมีความคิดนอกเนกนัย
4. การคิดแบบเอกนัย (Convergent Production) เป็นความสามารถในการคิดหาคำตอบ ที่ดีที่สุดหาเกณฑ์ที่เหมาะสมได้ดีที่สุด ดังนั้นคำตอบแบบนี้ก็ต้องถูกเพียงคำตอบเดียว
5. การคิดแบบประเมินค่า (Evaluation) เป็นความสามารถในการตีราคาลงสรุปโดยอาศัยเกณฑ์ที่ดีที่สุด

มิติที่ 2 ด้านเนื้อหา (Content) เป็นด้านที่ประกอบด้วยสิ่งเร้าและข้อมูลต่างๆแบ่งออกได้ 4 อย่าง คือ

1. ภาพ (Figural) หมายถึง สิ่งเร้าที่เป็นรูปธรรมหรือรูปที่แน่นอนสามารถจับต้องได้ หรือเป็นรูปภาพที่ระลึกนึกออกได้ดังรูปนั้นก็

2. สัญลักษณ์ (Symbolic) หมายถึง ข้อมูลที่เป็นเครื่องหมายต่างๆ เช่นตัวอักษร โน้ตดนตรี รวมทั้งสัญญาณต่างๆ ด้วย

3. ภาษา (Semantic) หมายถึง ข้อมูลที่เป็นถ้อยคำพูดหรือภาษาเขียนที่มีความหมายสามารถใช้ติดต่อสื่อสารแต่ละกลุ่มได้ แต่ส่วนใหญ่มองในด้านคิด (Verbal Thinking) มากกว่าการเขียน คือ มองความหมาย

4. พฤติกรรม (Behavioral) หมายถึง ข้อมูลที่เป็นการแสดงออก รวมทั้งทัศนคติ ความต้องการ การรับรู้ ความคิด

มิติที่ 3 ผลของการคิด (Products)เป็นผลของกระบวนการจัดกระทำของความคิดกับข้อเนื้อหา ผลผลิตของความคิดแยกได้เป็นรูปร่างต่างๆ กัน ซึ่งแบ่งออกได้ 6 อย่าง คือ

1. หน่วย (Units) หมายถึง สิ่งที่มีคุณสมบัติเฉพาะตัวและแตกต่างไปจากสิ่งอื่นๆ เช่น คน สุนัข แมว

2. จำพวก (Classes) หมายถึง ชุดของหน่วยที่มีคุณสมบัติร่วมกัน เช่น ข้าวโพด มะพร้าว เป็นพืชใบเลี้ยงเดี่ยวเหมือนกัน

3. ความสัมพันธ์ (Relation) หมายถึง ผลของการโยงความคิดสองประเภทหรือหลายประเภทเข้าด้วยกัน โดยอาศัยลักษณะบางประการเป็นเกณฑ์ อาจจะเป็นหน่วยกับหน่วย จำพวกกับจำพวก ระบบกับระบบ เช่น คนกับอาหาร ต้นไม้กับปุ๋ย

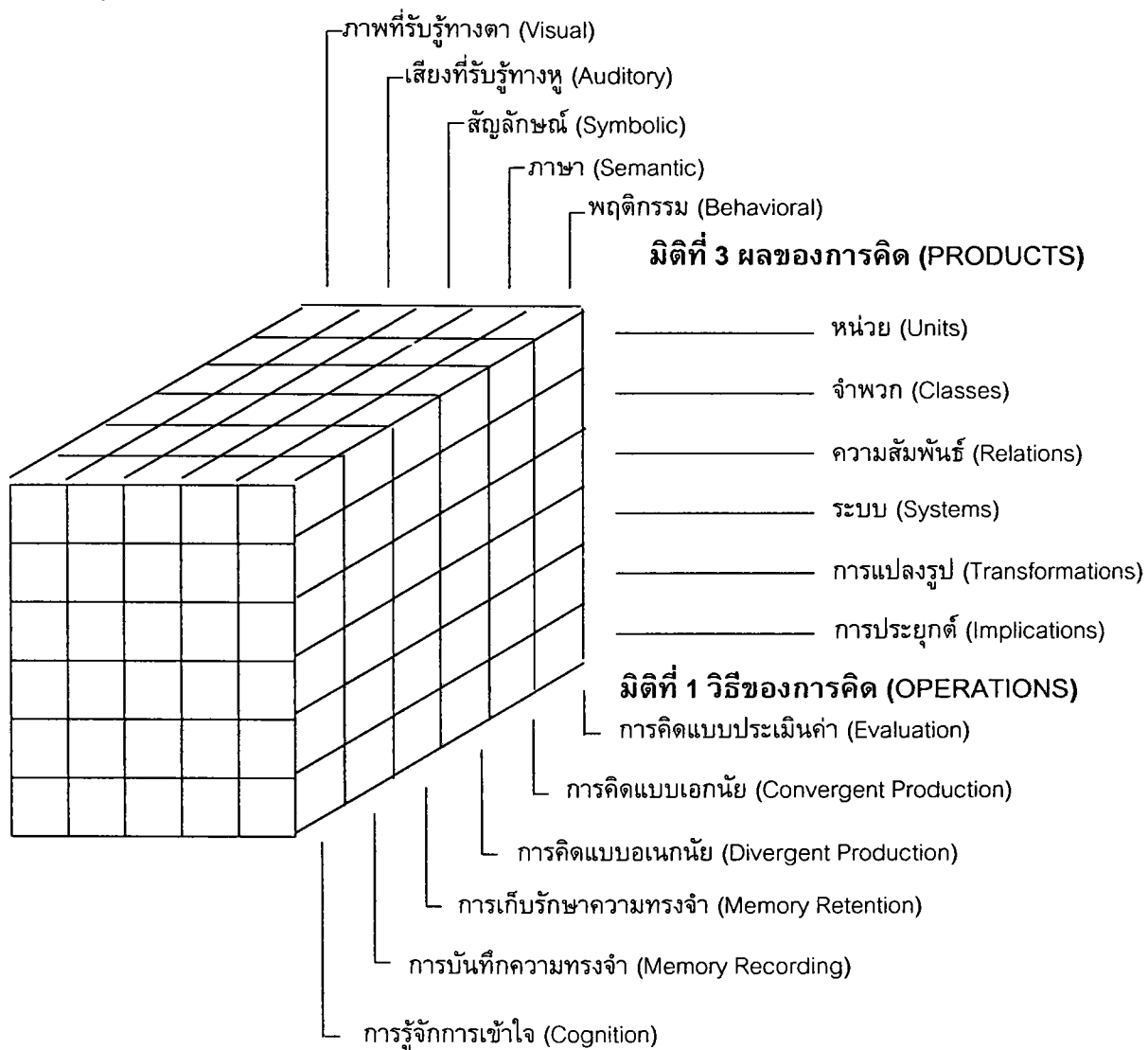
4. ระบบ (Systems) หมายถึง การจัดองค์การ จัดแบบแผนหรือจัดรวมโครงสร้างให้อยู่ในระบบว่าอะไรมาก่อนมาหลัง

5. การแปลงรูป (Transformation) หมายถึง การเปลี่ยนแปลงสิ่งที่มีอยู่ให้มีรูปแบบใหม่ การเปลี่ยนแปลงอาจจะมองในรูปแบบของข้อมูลหรือประโยชน์ก็ได้

6. การประยุกต์ (Implication) หมายถึง ความเข้าใจในการนำข้อมูลไปใช้ขยายความเพื่อการพยากรณ์หรือคาดคะเนข้อความในตรรกวิทยา ประเภท “ ถ้า...แล้ว...” ก็เป็นพวกใช้การคาดคะเนโดยอาศัยเหตุผล

เมื่อรวมทั้งสามมิติประกอบกัน จะเห็นว่ามีโครงสร้างการวัดเชาว์ปัญญาตามทฤษฎีของ กิลฟอร์ด ประกอบด้วย $5 \times 4 \times 6 = 120$ หน่วยลูกบาศก์ ต่อมาในปี ค.ศ.1977 กิลฟอร์ดได้เปลี่ยนแปลงเพิ่มเติมมิติที่ 2 ด้านเนื้อหา (Content) ในส่วนของภาพ (Figural) ออกเป็นภาพที่รับรู้ทางตา (Visual) และเสียงที่รับรู้ทางหู (Auditory) จึงทำให้มิติที่ 2 ด้าน เนื้อหาเพิ่มเป็น 5 ลักษณะ และโครงสร้างการวัดเชาว์ปัญญาตามทฤษฎีของกิลฟอร์ดก็เพิ่มขึ้นเป็น $5 \times 5 \times 6 = 150$ หน่วยลูกบาศก์ และในปีค.ศ.1988 กิลฟอร์ดก็เปลี่ยนแปลงองค์ประกอบในมิติที่ 1 ด้านกระบวนการ หรือวิธีการของการคิด (Operation) โดยขยายองค์ประกอบด้านความจำ (Memory Recording) จึงทำให้โครงสร้างการวัดเชาว์ปัญญาตามทฤษฎีของกิลฟอร์ดเปลี่ยนไปเป็น $6 \times 5 \times 6 = 180$ หน่วยลูกบาศก์ แบบจุลภาค (Micro-model) โดยในแต่ละตัวจะประกอบด้วย หน่วยย่อยของ 3 มิติ โดยเรียงจากวิธีของการคิด-เนื้อหา-ผลของการคิด (Operation-Content-Product)

มิติที่ 2 ด้านเนื้อหา (CONTENTS)



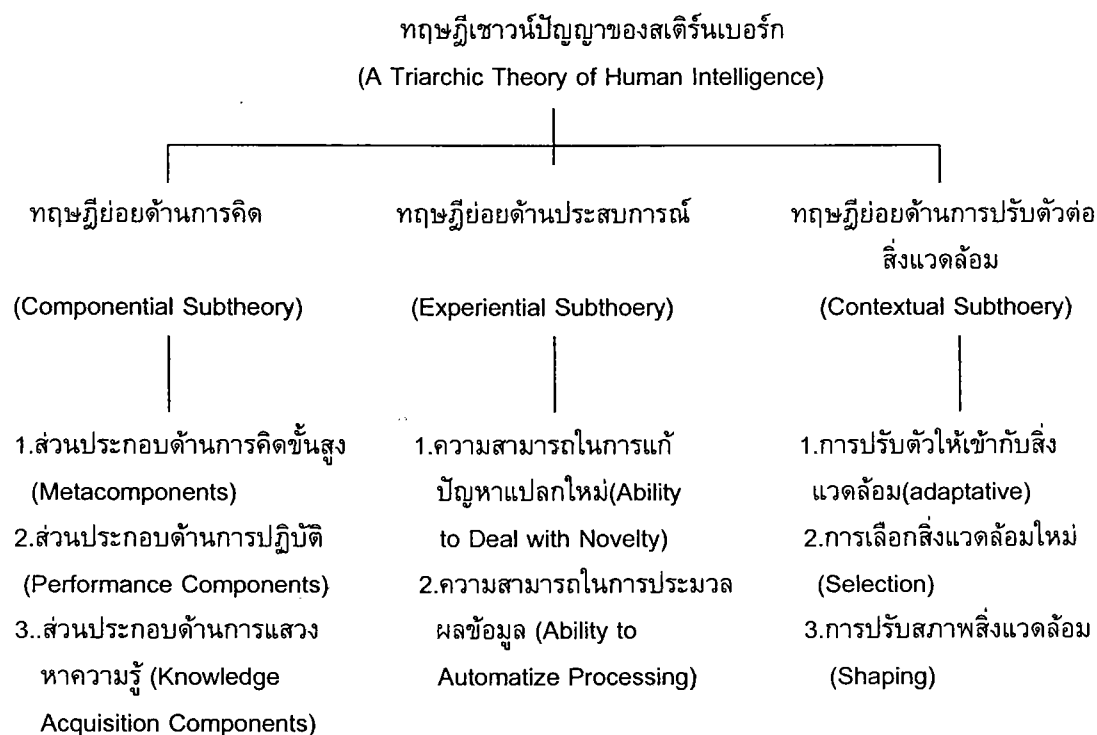
ภาพประกอบ 3 โครงสร้างความสามารถทางสมองของทฤษฎีโครงสร้างสามมิติของปัญญา

2.2.6. ทฤษฎีความสามารถทางสมองสองระดับ (Two-level Mental Ability) เป็นทฤษฎีที่เจนเซน (Jansen) ได้สร้างขึ้นในปี ค.ศ.1968 โดยกล่าวว่า ความสามารถทางสมองมีอยู่ 2 ระดับได้แก่ ระดับที่ 1 และระดับที่ 2 ดังนี้ (ชาญวิทย์ เทียมบุญประเสริฐ.2528:39-40)

ระดับที่ 1 เป็นความสามารถด้านการเรียนรู้และการจำ ที่เป็นการสะสมข้อมูลไว้พร้อมที่จะระลึกถึงข้อมูลนั้นได้ แต่จะไม่มีการแปลงรูปหรือมีการคิดโดยการผ่านกระบวนการทางสมองแต่อย่างใด

ระดับที่ 2 เป็นความสามารถทางการจัดการกระทำทางสมองที่เป็นการสร้างมโนภาพ คิดหาเหตุผล ตลอดจนการแก้ปัญหาต่างๆ ซึ่งเปรียบได้กับองค์ประกอบทั่วไป (G-Factor)

2.2.7. ทฤษฎีเชาว์ปัญญาสามหลัก (A Triarchic Theory of Human Intelligence) เป็นทฤษฎีที่คิดขึ้นโดยสเติร์นเบอร์ก เมื่อปี ค.ศ.1985 ที่กล่าวว่าเชาว์ปัญญาประกอบด้วย 3 ทฤษฎีย่อย คือ ทฤษฎีย่อยด้านการคิด (Componential Subtheory), ทฤษฎีย่อยด้านประสบการณ์ (Experiential Subtheory) และ ทฤษฎีย่อยด้านการปรับตัวต่อสิ่งแวดล้อม (Contextual Subtheory) ดังภาพประกอบต่อไปนี้ (พิสมัย สารกุล.2542:26-36; อ้างอิงจากStrenberg.1985:320-321)



ภาพประกอบ 4 โครงสร้างทฤษฎีเชาว์ปัญญาของสเติร์นเบอร์ก

2.2.7.1. ทฤษฎีย่อยด้านการคิด (Componential Subtheory) เป็นกระบวนการทางสมองที่เป็นพื้นฐานสำคัญในการแก้ปัญหา และเป็นกระบวนการประมวลผลข้อมูลเบื้องต้น สามารถแบ่งรูปแบบตามหน้าที่พื้นฐานได้ 3 ลักษณะ คือ ส่วนประกอบด้านการคิดขั้นสูง(Metacomponent) ส่วนประกอบด้านการปฏิบัติการ(Performance components) และส่วนประกอบด้านการแสวงหาความรู้ (Knowledge acquisition component)

2.2.7.2. ทฤษฎีย่อยด้านประสบการณ์ (Experiential Subtheory) เป็นความสามารถในการแก้ปัญหาแปลกใหม่ และความสามารถในการประมวลผลข้อมูล

2.2.7.3. ทฤษฎีย่อยด้านการปรับตัวต่อสิ่งแวดล้อม (Contextual Subtheory) เป็นกิจกรรมทางสมองที่บุคคลแสดงการกระทำโดยตรงต่อสิ่งแวดล้อมเพื่อการปรับตัวเข้ากับสิ่งแวดล้อม เพื่อเลือกสิ่งแวดล้อมใหม่ และเพื่อปรับสภาพแวดล้อมให้เหมาะสม กับการดำรงชีวิตของบุคคล

2.2.8. ทฤษฎีเชาว์ปัญญาหลากหลาย (Theory of Multiple Intelligence) เป็นทฤษฎีที่นำเสนอโดย การ์ดเนอร์ (Gardner) ในปีค.ศ.1983 ซึ่งได้นิยามว่าเชาว์ปัญญาเป็นวิสัยความสามารถในการแก้ปัญหา หรือสร้างผลงานที่มีค่าในกลุ่มวัฒนธรรมต่างๆ รวมเป็นความสามารถทั้งหมด 8 ด้าน คือ (อารี สันทนต์.25 35:1-2)

2.2.8.1. ปัญญาด้านภาษา (Linguistic Intelligence) เป็นความสามารถสูงในการใช้ภาษาพูด ภาษาเขียน การจัดกระทำเกี่ยวกับโครงสร้างของภาษา เสียง อธิบายความหมาย และเรื่องเกี่ยวกับภาษา

2.2.8.2. ปัญญาด้านตรรกและคณิตศาสตร์ (Logical-Mathematical Intelligence) เป็นความสามารถสูงในการใช้ตัวเลข แล้วยังรวมถึงความไวในการเห็นความสัมพันธ์ แบบแผน ตรรกวิทยา การคิดเชิงเหตุผล การคิดเชิงนามธรรม และการคิดที่เป็นเหตุเป็นผล

2.2.8.3. ปัญญาทางด้านมิติ (Spatial Intelligence) เป็นความสามารถสูงในการมองเห็นพื้นที่ และสามารถปรับปรุงและคิดวิธีการใช้เนื้อที่ได้ดี ความสามารถที่จะมองเห็น และแสดงออกเป็นรูปร่างถึงสิ่งที่เห็นและความคิดเกี่ยวกับพื้นที่

2.2.8.4 ปัญญาทางด้านร่างกายและการเคลื่อนไหว (Bodily – Kinesthetic Intelligence) เป็นความสามารถสูงในการแสดงความสามารถสูงในการใช้ร่างกายของตนแสดงความคิด ความรู้สึก และความสามารถในการใช้มือประดิษฐ์ ซึ่งปัญญาทางด้านนี้รวมถึงทักษะทางร่างกายด้วย

2.2.8.5 ปัญญาทางด้านดนตรี (Musical Intelligence) เป็นความสามารถสูงทางด้านดนตรี เช่น นักดนตรี นักแต่งเพลง รวมถึงความไวในเรื่องจังหวะ ทำนองเสียงความสามารถในการเข้าใจและวิเคราะห์ดนตรี

2.2.8.6 ปัญญาทางด้านมนุษยสัมพันธ์ (Interpersonal Intelligence) เป็นความสามารถสูงในการเข้าใจ อารมณ์ ความรู้สึก ความคิดและเจตนาของผู้อื่น รวมถึงความไวในการสังเกต น้ำเสียง ใบหน้า ท่าทาง ทั้งยังมีความสามารถสูงในการรู้ถึงลักษณะต่างๆ ของสัมพันธ์ภาพของมนุษย์ และสามารถตอบสนองได้เหมาะสมและมีประสิทธิภาพ เช่น สามารถทำให้บุคคลหรือกลุ่มบุคคลปฏิบัติตาม

2.2.8.7 ปัญญาทางด้านตนหรือการเข้าใจตัวเอง (Intrapersonal Intelligence) เป็นความสามารถสูงในการรู้จักตนเอง และประพฤติ ปฏิบัติตนได้จากความรู้จักนี้ ได้แก่ รู้จักตัวเองตามความเป็นจริงเช่น มีจุดอ่อน จุดแข็งเรื่องใด รู้เท่าทันอารมณ์ ความคิด ความต้องการของคน และสามารถที่จะฝึกตนและเข้าใจตนเอง

2.2.8.8 ปัญญาทางด้านธรรมชาติวิทยา (Naturalist Intelligence) เป็นความสามารถในการรู้จักธรรมชาติของพืชและสัตว์ สามารถจำแนกประเภท

จากการศึกษาทฤษฎีเกี่ยวกับความถนัดต่างๆ ดังที่กล่าวมานี้พบว่าทฤษฎีส่วนใหญ่จะมุ่งอธิบายถึงความสามารถทางด้านสติปัญญา ที่มีองค์ประกอบสำคัญ คือ ความสามารถทางภาษา และความ

สามารถทางตัวเลข ซึ่งสอดคล้องกับแบบทดสอบวัดความถนัด ที่ผู้วิจัยจะได้ทำการสร้างตามแนวคิดตาม ทฤษฎีหลายองค์ประกอบของเทอร์สโตน โดยแบ่งออกเป็น 3 ด้าน คือ องค์ประกอบด้านตัวเลข องค์ประกอบ ด้านเหตุผล และองค์ประกอบด้านภาษา เพื่อเป็นประโยชน์ในการศึกษาค้นคว้าเกี่ยวกับการทำหน้าที่ต่างกัน ของข้อสอบแล้ว ยังสามารถประเมินความสามารถของบุคคลต่างๆ ทั้งความสามารถโดยรวมและความ สามารถเฉพาะกิจกรรมได้ถูกต้อง

3. งานวิจัยที่เกี่ยวข้อง

3.1 งานวิจัยต่างประเทศ

สำหรับงานวิจัยต่างประเทศที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบนั้นได้มีผู้วิจัยไว้หลาย ท่านด้วยกัน ซึ่งผู้วิจัยได้รวบรวมและสรุปไว้ ดังนี้

ผจงจิต อินทรสุวรรณ (Intrasuwan.1979:92) ได้ศึกษาวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบที่ ต่างกัน 3 วิธี คือ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ วิธีราสซัส วิธีไคสแควร์ ของแบบทดสอบ IEA (The International Association for Evaluation Achievement) กลุ่มตัวอย่างที่ใช้ศึกษาเป็นนักเรียนอายุ 14 ปี ชาวอังกฤษ อเมริกัน และนิวซีแลนด์ จำนวน 8,601 คน ผลการวิเคราะห์พบว่า สหสัมพันธ์ระหว่างวิธี ทั้ง 3 วิธี มีค่าอยู่ระหว่าง 0.051-0.98 และวิธีไคสแควร์กับวิธีราสซัส มีสหสัมพันธ์กันสูงถึง 0.98

ดูลิทเทิล และ เคลียร์ (Doolittle and Cleary.1987:157-166) ได้ศึกษาตรวจค้นการทำหน้าที่ต่างกัน ระหว่างเพศของข้อสอบในแบบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ โดยใช้กลุ่มตัวอย่างจำนวน 8 กลุ่ม ที่ทำแบบ ทดสอบ ACT (Assessment Mathematics Test) รวม 8 ฉบับ กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นเพศชายและ เพศหญิงจำนวน 1,300-1,400 คน โดยเป็นหญิงประมาณร้อยละ 55 วิเคราะห์ด้วยดัชนีของ ลินน์ และ ฮาร์นิสซ์ (Linn and Harnisch.1981) ซึ่งเป็นโมเดลโลจิสติกแบบ 3 พารามิเตอร์ ถ้าดัชนี Z ที่ได้เป็นบวก แสดงว่าง่าย สำหรับกลุ่มเปรียบเทียบถ้าเป็นลบแสดงว่ายากสำหรับกลุ่มเปรียบเทียบ ผลการวิเคราะห์พบว่า ดัชนี Z มีค่าเป็นลบ สำหรับข้อสอบที่วัดด้านเรขาคณิต และการใช้เหตุผลเชิงพีชคณิต และเลขคณิตในทุก ฉบับแสดงว่ายากสำหรับชายมากกว่าหญิง และผลจากการวิเคราะห์ด้วยวิธีวิเคราะห์ความแปรปรวน ได้ผล เช่นเดียวกับการคำนวณค่า Z

ทิสเซน สเตนเบอร์ก และ ไวน์เนอร์ (Tissen, Steinberg and Wainer.1988:147-170) ได้ศึกษาความ แตกต่างของโค้งการตอบข้อสอบระหว่างกลุ่มด้วยทฤษฎีการตอบข้อคำถาม เพื่อศึกษาถึงสภาพการณ์ในการ ทดสอบสมมติฐานความเท่ากันของค่าพารามิเตอร์ ด้วยค่าไคสแควร์แบบอัตราส่วน Likelihood (IRT-LR) และ วิธีแมนเทิล-แฮนส์เซล โดยการจำลองข้อมูล

ผลการศึกษาพบว่า วิธีไคสแควร์แบบอัตราส่วน Likelihood (IRT-LR) สามารถใช้วิเคราะห์การทำ หน้าที่ต่างกันของข้อสอบได้เมื่อมีจำนวนกลุ่มตัวอย่างน้อยๆ คือ ประมาณ 500 คน ความแตกต่างของ ค่าความยากของข้อสอบระหว่างกลุ่มมีค่าเท่ากับ 0.3 และความแตกต่างในด้านค่าสัมประสิทธิ์การเดามีค่าเท่า กับ 0.1 โดยจำนวนข้อสอบรวมจะมากหรือน้อยก็ไม่มีผลกระทบ ถ้าความแตกต่างมีค่าน้อยกว่านี้ไม่สามารถ ที่จะใช้วิธีไคสแควร์แบบอัตราส่วน Likelihood (IRT-LR) วิเคราะห์ข้อสอบที่ทำหน้าที่ต่างกันได้ แต่ถ้าใช้ กลุ่มตัวอย่างมากกว่านี้ก็จะทำการวิเคราะห์ได้ การระบุถึงการทำหน้าที่ต่างกันของข้อสอบใช้วิธีเปรียบเทียบความแตกต่างระหว่างกลุ่มในข้อสอบที่ต้องการวิเคราะห์การทำหน้าที่ต่างกันกับข้อสอบรวม ซึ่งมีข้อ ตกกลางว่าข้อสอบรวมนั้นทำหน้าที่ไม่ต่างกัน สำหรับผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธี ไคสแควร์แบบอัตราส่วน Likelihood (IRT-LR) กับวิธีของแมนเทิล-แฮนส์เซล ทั้งสองวิธีให้ผลการวิเคราะห์

คล้ายคลึงกัน ข้อเสนอแนะจากการวิจัยครั้งนี้ คือ วิธีของแมนเทล-แฮนด์เซล อาจจะใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของแบบทดสอบยาวๆ หรืออาจใช้เป็นเครื่องมือในการกลั่นกรองข้อสอบก่อนที่จะใช้วิธีไคสแควร์แบบอัตราส่วน Likelihood (IRT-LR) ต่อไป

แฮมเบลตัน และ โรเจอร์ (Hambleton and Roger.1989:313-334) ได้ศึกษาการเปรียบเทียบวิธีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ 2 วิธี ได้แก่ วิธีแมนเทล-แฮนด์เซล และ วิธีทฤษฎีการตอบข้อสอบ โดยการใช้การวัดพื้นที่ระหว่างโค้งลักษณะข้อสอบ โดยใช้ข้อมูลจากการตอบแบบทดสอบ NMHSPE (New Mexico High School Proficiency Exam) ซึ่งเป็นแบบทดสอบวัดทักษะการดำเนินชีวิต 5 ด้าน ดังนี้ ด้านที่ 1 วัดความรู้เกี่ยวกับแหล่งทรัพยากรชุมชน ด้านที่ 2 วัดความรู้ผู้บริโภค ด้านที่ 3 วัดความรู้เกี่ยวกับรัฐบาลและกฎหมาย ด้านที่ 4 วัดความรู้ด้านสุขภาพกายและจิต และด้านที่ 5 วัดความรู้ด้านอาชีพ ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 150 ข้อ โดยที่ผู้วิจัยได้ตัดข้อสอบที่ง่ายมาก ($p > 0.70$) และค่าอำนาจจำแนกต่ำ ($r < 0.10$) ได้ข้อสอบ 75 ข้อ วิธีแมนเทล-แฮนด์เซล วิเคราะห์โดยใช้ 76 ระดับคะแนน ส่วนวิธีทฤษฎีการตอบข้อสอบใช้การวัดพื้นที่ระหว่างโค้งลักษณะข้อสอบ (ICC) มีพิสัยของความสามารถที่นำมาใช้ในการคำนวณพื้นที่ที่กำหนด 2 ค่า เบี่ยงเบนมาตรฐาน สูงกว่าความสามารถเฉลี่ยของนักเรียนชาวอเมริกันผิวขาว และ 2 ค่าเบี่ยงเบนมาตรฐาน สูงกว่าความสามารถเฉลี่ยของชาวอเมริกันพื้นเมือง กลุ่มตัวอย่างที่ใช้ในการศึกษาได้แก่ นักเรียนชาวอเมริกันผิวขาวและอเมริกันพื้นเมืองที่สุ่มมาอย่างละ 2,000 คน โดยแบ่งออกเป็น 2 กลุ่มๆ ละ 1,000 คน ตามลำดับคู่ คือ

ผลการศึกษาพบว่า วิธีแมนเทล-แฮนด์เซลมีความสอดคล้องของการตรวจสอบร้อยละ 80 ส่วนวิธีโค้งลักษณะข้อสอบ ร้อยละ 73 และพบว่า ร้อยละ 47 ของข้อสอบพบการทำหน้าที่ต่างกันในการวิเคราะห์คู่ที่ 1 เป็นข้อสอบที่ทำหน้าที่ต่างกันในการวิเคราะห์คู่ที่ 2 ด้วยในวิธีแมนเทล-แฮนด์เซล และร้อยละ 61 ในวิธีโค้งลักษณะข้อสอบ ในทางตรงข้ามพบว่า ร้อยละ 64 ของข้อสอบที่ทำหน้าที่ต่างกันในการวิเคราะห์คู่ที่ 2 เป็นข้อสอบที่ทำหน้าที่ต่างกันในการวิเคราะห์คู่ที่ 1 ด้านในวิธีแมนเทล-แฮนด์เซล และร้อยละ 56 ในวิธีโค้งลักษณะข้อสอบจากผลการศึกษาดังกล่าวจะพบว่า ความคงที่ของการวิเคราะห์ทั้งวิธีแมนเทล-แฮนด์เซล และวิธีโค้งลักษณะข้อสอบ อยู่ในระดับปานกลาง อาจเป็นเพราะกลุ่มตัวอย่างที่ใช้ค่อนข้างน้อย

สวามินาธาน และโรเจอร์ (Swaminathan and Roger. 1990:361-370) ได้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีถดถอยโลจิสติก (LR-DIF) กับวิธีแมนเทล-แฮนด์เซล (MH) โดยใช้ขนาดตัวอย่าง 2 ขนาด คือ 200 และ 500 คน และความยาวของแบบทดสอบ 3 ขนาด คือ 40, 60 และ 80 ข้อ โดยจำลองข้อมูลการตอบข้อสอบด้วยโปรแกรม DATAGEN แบบ 3 พารามิเตอร์ และข้อมูลการจำลองการทำหน้าที่ต่างกัน ใช้แบบ 2 พารามิเตอร์ จำลองข้อมูลที่ทำหน้าที่ต่างกัน 2 แบบ คือ การทำหน้าที่ต่างกันของข้อสอบแบบยูนิฟอร์ม (Uniform DIF) และการทำหน้าที่ต่างกันของข้อสอบแบบนอยูนิฟอร์ม (Nonuniform DIF) โดยให้ค่าจำแนกของทั้ง 2 กลุ่มเท่ากัน ส่วนค่าความยากของข้อสอบแปรเปลี่ยนไปตามกลุ่มเพื่อให้เกิดระดับการทำหน้าที่ต่างกันต่างๆ กันตามต้องการ แต่ในลักษณะการทำหน้าที่ต่างกันของข้อสอบแบบนอยูนิฟอร์ม (Nonuniform DIF) นั้นให้ค่าความยากของข้อสอบของ 2 กลุ่มเท่ากัน แต่ค่าอำนาจจำแนกแปรเปลี่ยนไปตามกลุ่ม

ผลการศึกษา พบว่าการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทั้ง 2 วิธี ให้ผลใกล้เคียงกัน ในกรณีการทำหน้าที่ต่างกันของข้อสอบแบบยูนิฟอร์ม (Uniform DIF) วิธีแมนเทล-แฮนด์เซลตรวจพบว่าการทำหน้าที่ต่างกันของข้อสอบได้ดีกว่าวิธีถดถอยโลจิสติกเล็กน้อย โดยมีการตรวจพบการทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 75 กรณีใช้กลุ่มตัวอย่าง 200 คน และตรวจพบได้ถูกต้องร้อยละ 100 ในกรณีใช้กลุ่มตัวอย่าง

500 คน ในทุกความยาวของแบบทดสอบ แต่ในกรณีของการทำหน้าที่ต่างกันของข้อสอบแบบนูนนิฟอร์ม (Nonuniform DIF) นั้น วิธีแมนเทล-แฮนดเซล ไม่สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ ส่วนวิธีการถดถอยโลจิสติก ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้องร้อยละ 50 ในกรณีกลุ่มตัวอย่างน้อย และแบบทดสอบสั้น และตรวจสอบการทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 75 ในกรณีกลุ่มตัวอย่างขนาดใหญ่และแบบทดสอบยาว สำหรับการตรวจสอบระดับความคลาดเคลื่อนประเภทที่ 1 (Type 1 error rate) พบว่าวิธีแมนเทล-แฮนดเซล มีประสิทธิภาพดีกว่า โดยตรวจสอบผิดพลาดประมาณร้อยละ 1 ส่วนวิธีการถดถอยโลจิสติก ตรวจสอบผิดพลาดประมาณร้อยละ 1-6 นอกจากนี้ยังพบว่าวิธีถดถอยโลจิสติก จะเสียค่าใช้จ่ายมากกว่าวิธีแมนเทล-แฮนดเซล ประมาณ 3-4 เท่า

ราชู ดาสโก และสเลนเด้ (Raju, Drasgow and Slinde.1993:301-314) ได้ทำการศึกษาเปรียบเทียบผลการทดสอบค่า Z ของราชู ที่ได้จากการพื้นที่ชนิดมีและไม่มีเครื่องหมาย การทดสอบค่า χ^2 ของ Lord (ใช้โมเดล 2 พารามิเตอร์) และการทดสอบ MH- χ^2 ข้อมูลที่ใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบคือ ผลการตอบแบบทดสอบ GMRT (Gater-Mac Ginitie Reading Test) เฉพาะแบบทดสอบย่อยที่วัดคำศัพท์ 45 ข้อ เป็นชนิดเลือกตอบ 5ตัวเลือก กับนักเรียนระดับ 10 และ 12 ปีการศึกษา 1987 จำนวน 839 คนเป็นนักเรียนผิวดำ 245 คน ผิวดำ 436 คน เป็นนักเรียนหญิง 440 คน และนักเรียนชาย 399 คน

ผลการศึกษาพบว่าข้อสอบที่ทำหน้าที่ต่างกัน 1 ข้อ สำหรับการทดสอบด้วย χ^2 และ Z ทั้งพื้นที่ชนิดคิดเครื่องหมายและไม่คิดเครื่องหมายได้แก่ ข้อ 41 ส่วนวิธีแมนเทล-แฮนดเซลพบข้อสอบที่ทำหน้าที่ต่างกัน 2 ข้อ ได้แก่ ข้อ 14 และ 27 ซึ่งทั้ง 3 ข้อเป็นข้อสอบที่ยากต่างกันมากสำหรับนักเรียนผิวดำและผิวดำ

นารายานาน และ สวามินาธาน (Narayanan and Swaminathan.1994:315-328) ได้ศึกษาเปรียบเทียบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีแมนเทล-แฮนดเซล และวิธีซิปเทสต์ ในข้อสอบแบบนูนนิฟอร์ม โดยมีขนาดกลุ่มตัวอย่าง กลุ่มอ้างอิงมี 3 ขนาด ได้แก่ 300, 500 และ 1,000 คน กลุ่มเปรียบเทียบ 3 ขนาด ได้แก่ 100, 200 และ 300 คน ซึ่งจับคู่ศึกษาได้ 9 เงื่อนไข

ผลการศึกษาพบว่า วิธีแมนเทล-แฮนดเซล และวิธีซิปเทสต์ สามารถตรวจสอบการทำหน้าที่ต่างกันได้ดีในขนาดกลุ่มตัวอย่าง 600 คน (กลุ่มอ้างอิงและกลุ่มเปรียบเทียบกลุ่มละ 300 คน) ขนาดของการทำหน้าที่ต่างกันทั้ง 4 ขนาด มีผลต่อการตรวจสอบการทำหน้าที่ต่างกันทั้ง 2 วิธีอย่างมีนัยสำคัญทางสถิติ ความแตกต่างในการแจกแจงความสามารถไม่มีผลต่อวิเคราะห์การทำหน้าที่ต่างกันด้วยวิธีซิปเทสต์ แต่มีผลต่อวิธี แมนเทล-แฮนดเซลเป็นเพราะ วิธีซิปเทสต์ ได้มีการปรับแก้การถดถอยขนาดของแบบทดสอบ และร้อยละของข้อสอบที่ทำหน้าที่ต่างกันไม่มีผลต่อวิธีทั้งสอง สำหรับการเกิดความคลาดเคลื่อนแบบที่ 1 จะมีอัตราการเกิดอยู่ในระดับปกติ แต่วิธีซิปเทสต์จะเพิ่มขึ้นเล็กน้อยในกลุ่มที่มีความสามารถเท่ากัน แต่ในกลุ่มที่มีความสามารถเท่ากัน แต่ในกลุ่มที่มีความสามารถไม่เท่ากันเกิดความคลาดเคลื่อนแบบที่ 1 จะมีอัตราการเกิดสูงขึ้นทั้งสองวิธี ซึ่งเป็นการเพิ่มความยากของแบบทดสอบจะทำให้เกิดความเชื่อมั่นสูงขึ้น ซึ่งมีผลให้เกิดความเชื่อมั่นสูงขึ้น ซึ่งมีผลให้ความคลาดเคลื่อนแบบที่ 1 น่าจะมีอัตราการเกิดลดลง

ซาง มัสซิโอ และรอลอส (Chang Mazzeo and Roussos.1996:333-353) ได้ทำการศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของแบบทดสอบที่ตรวจให้คะแนนแบบหลายค่า (Polytomously) โดยวิธี SIBTEST โดยใช้ข้อมูลจำลอง มีวัตถุประสงค์เพื่อใช้เปรียบเทียบวิธี Modified SIBTEST กับ วิธี แมนเทล-แฮนดเซล (Mantel-Haenszel) และวิธี Standardized Mean Difference (SMD) ทำการศึกษา 2 ครั้ง ดังนี้ การศึกษาครั้งที่ 1 ทำการศึกษาเปรียบเทียบวิธีทั้ง 2 วิธีการ ภายใต้เงื่อนไขของการศึกษาแบบแมนเทล-แฮนดเซล และแบบ SMD จะมีการศึกษาการทำหน้าที่ต่างกันของข้อสอบว่าดีหรือไม่

ผลการศึกษาพบว่า วิธี Modified SIBTEST ใช้ในทางปฏิบัติจะเหมาะสมมากกว่าวิธีอื่นแต่วิธีแมนเทล-แฮนด์เซล และวิธี SMD ใช้ได้ดีกับข้อมูลหาย การศึกษาครั้งที่ 2 ใช้ข้อมูลภายใต้เงื่อนไขวิธีการทดสอบการทำหน้าที่ต่างกันของคะแนนที่สังเกตได้ของแบบทดสอบที่มีการตรวจให้คะแนนแบบ 0-1 (Dichotomous) ที่ไม่สมบูรณ์ ผลลัพธ์ของการศึกษาครั้งนี้พบว่า ภายใต้เงื่อนไขทั้งหมด วิธี Modified SIBTEST ใช้ได้ดีภายใต้ความคลาดเคลื่อน แบบ 1 มากกว่ากระบวนการอื่น และวิธีการอื่น

อัลลาลอฟ แฮมเบิลตัน และสิริซี (Allalovf Hambleton and Sireci.1999:185-198) ได้ทำการศึกษาดึงสาเหตุของการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบแปลความทางภาษาโดยเสนอรูปแบบของข้อสอบที่ทำหน้าที่ต่างกัน และสาเหตุของการทำหน้าที่ต่างกันในรูปแบบแผนการเขียนข้อสอบ โดยมีวัตถุประสงค์เพื่อศึกษาถึงสาเหตุของการทำหน้าที่ต่างกันของข้อสอบแปลความทางภาษา โดยแบ่งเป็น 2 ส่วน ส่วนแรกศึกษาเกี่ยวกับรูปแบบของข้อสอบทางภาษากับการทำหน้าที่ต่างกันของข้อสอบ เมื่อมีการแปลความจากภาษาหนึ่งไปยังอีกภาษาหนึ่ง ส่วนที่สองเกี่ยวกับการค้นหาสาเหตุของการทำหน้าที่ต่างกันของข้อสอบโดยวิธีวิเคราะห์ 2 วิธี คือ วิธีการวิเคราะห์การทำหน้าที่ต่างกันโดยดูจากค่าสถิติ และวิธีการวิเคราะห์โดย นักแปลภาษาในด้านรูปแบบและเนื้อหา จากข้อสอบที่ทำหน้าที่ต่างกัน เครื่องมือที่ใช้เป็นแบบทดสอบแปลความทางภาษาจากภาษาฮีบรู (Hebrew) เป็นภาษารัสเซีย ซึ่งมีด้านย่อยๆ 4 ด้าน คือ ด้านเหตุผล จำนวน 26 ข้อ ด้านความเข้าใจประโยค จำนวน 33 ข้อ ด้านตรรกศาสตร์ จำนวน 36 ข้อ และด้านความเข้าใจในการอ่าน จำนวน 30 ข้อ รวม 125 ข้อ ซึ่งใช้วิธีการวิเคราะห์ของแมนเทล-แฮนด์เซล ที่คำนวณจากโปรแกรม DICHODIF

ผลการศึกษาพบว่าข้อสอบแปลความทางภาษา 34% มีการทำหน้าที่ต่างกันของข้อสอบในด้านภาษา โดยแบบทดสอบด้านเหตุผลจะแสดงการทำหน้าที่ต่างกันของข้อสอบถึง 65 % ในแบบฝึกหัดการพูดภาษารัสเซียข้อสอบความเข้าใจประโยคจะมีการทำหน้าที่ต่างกันถึง 45% และเหตุผลหลักที่ข้อสอบทำหน้าที่ต่างกันซึ่งระบุจากนักแปลภาษาจำนวน 5 คน และนักวิจัยเกี่ยวกับการพูดภาษาฮีบรูจำนวน 3 คน รวม 8 คน พบว่าเกิดจากการเปลี่ยนความยากของคำ การเปลี่ยนรูปแบบของข้อสอบ ความแตกต่างของวัฒนธรรม และการเปลี่ยนในด้านเนื้อหา

สติเกอร์ และเอเมอร์ริช (Stricker and Emmerrich.1999:347-366) ได้ทำการศึกษาการทำหน้าที่ต่างกันของแบบทดสอบทางจิตวิทยา (The Advance Placement Psychology Examination : AP Test) ระหว่างความคุ้นเคย ความสนใจ และความพอใจ ที่แบ่งตามเนื้อหาวิชา จำนวน 100 ข้อ มีกลุ่มตัวอย่างเป็นนักเรียนหลักสูตรจิตวิทยา จำนวน 717 คน จำแนกเป็นเพศชาย 265 คน เพศหญิง 452 คน โดยวิธีแมนเทล-แฮนด์เซล และวิธีเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยมาตรฐาน พบว่าวิธีเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยมาตรฐานในเพศที่ต่างกัน เพศหญิงจะได้เปรียบในหัวข้อที่เกี่ยวกับแรงจูงใจ เพศชาย จะได้เปรียบในหัวข้อที่เกี่ยวกับความรู้สึกและการรับรู้ ในด้านความคุ้นเคย ความพอใจ ในเนื้อหาของข้อสอบที่ทำการทดสอบมีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ด้านความสนใจ ในแต่ละหัวข้อขึ้นอยู่กับเนื้อหาของข้อสอบที่ทำการทดสอบได้แก่ด้าน แรงจูงใจ จิตวิทยาพัฒนาการ ในเพศหญิงจะมีระดับความสนใจมากกว่าเพศชาย และมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 วิธีแมนเทล-แฮนด์เซล พบว่ามีค่าความแตกต่างในด้านความคุ้นเคย ความสนใจ และความพอใจ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ การทำหน้าที่ต่างกัน กรณีเพศต่างกัน สำหรับแต่ละหัวข้อของข้อสอบมีความสัมพันธ์เชิงเส้น เพศหญิงมีความคุ้นเคย ความสนใจ ความพอใจในแต่ละหัวข้อที่ทำการทดสอบมากกว่าเพศชาย

3.2 งานวิจัยในประเทศ

สำหรับงานวิจัยในประเทศไทยที่เกี่ยวข้องกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ หรือการวิเคราะห์ความลำเอียงของข้อสอบ มีดังนี้

พัชรีย์ ปิยภัณฑ (2531:64-69) ได้ทำการวิจัยเรื่อง “การวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน” โดยมีกลุ่มตัวอย่างเป็นนักเรียนชายและหญิง ชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2530 ของโรงเรียนสังกัดคณะกรรมการการศึกษาแห่งชาติ ในจังหวัดกรุงเทพมหานคร และจังหวัดสมุทรสาคร จำนวน 1,422 คน ซึ่งเลือกโดยวิธีการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ คือ แบบทดสอบวิชาคณิตศาสตร์ เรื่องสมการ จำนวน 45 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง ระหว่างกลุ่มผู้สอบที่มีเพศ และสภาพภูมิศาสตร์ต่างกัน ด้วยวิธีวิเคราะห์ 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีไคสแควร์ และวิธีโค้งลักษณะข้อสอบ

ผลการวิจัยพบว่า จำนวนข้อสอบที่ลำเอียงจากวิธีวิเคราะห์ทั้ง 3 วิธี ต่างกัน โดยวิธีโค้งลักษณะข้อสอบ พบจำนวนข้อสอบที่ลำเอียงมากที่สุด ค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงต่อกลุ่มนักเรียนกรุงเทพมหานครกับกลุ่มนักเรียนสมุทรสาคร มีค่าระหว่าง .1868 ถึง .6009 สำหรับค่าความเชื่อมั่นของแบบทดสอบก่อนและหลังการคัดเลือกข้อสอบที่ลำเอียงออกแล้วพบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เจน จันทรแสง (2537:84-91) ได้ทำการวิจัยเรื่อง “การวิเคราะห์ความลำเอียงของข้อสอบของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ” โดยมีกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ในจังหวัดนครสวรรค์ จำนวน 1,600 คน ซึ่งเลือกมาโดยวิธี การสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ คือ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 จำนวน 40 ข้อ ด้วยวิธีการเปรียบเทียบวิธีวิเคราะห์ ความลำเอียง 4 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีวิเคราะห์ความแปรปรวน วิธีไคสแควร์ และวิธี โค้งลักษณะข้อสอบ 1 พารามิเตอร์ ศึกษาสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีความลำเอียงทั้ง 4 วิธี และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบก่อนและหลังการคัดเลือก

ผลการวิจัยพบว่า วิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์มีจำนวนข้อสอบที่ลำเอียงมากที่สุด รองลงมา คือ วิธีวิเคราะห์ความแปรปรวน และวิธีไคสแควร์ สำหรับแบบทดสอบที่คัดเลือกข้อสอบที่ลำเอียงออกไปแล้ว ค่าความเชื่อมั่นลดลงจากเดิมอย่างไม่มีนัยสำคัญทางสถิติและเมื่อปรับค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่ลำเอียงออกแล้วค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

นิรมล ชัยสวัสดิ (2537:57-58) ได้ทำการวิจัยเรื่อง “การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของแบบทดสอบความเข้าใจในการอ่านภาษาไทย ตามทฤษฎีคลาสสิกอลที่ใช้วิธีวิเคราะห์ต่างกัน” โดยศึกษาเปรียบเทียบผลการวิเคราะห์ความลำเอียงระหว่างกลุ่มผู้สอบที่มีเพศต่างกัน ด้วยวิธีวิเคราะห์ความลำเอียงของข้อสอบตามทฤษฎีคลาสสิกอล 3 วิธี คือ วิธีแปลงค่าความยาก วิธีไคสแควร์ และวิธีวิเคราะห์ความแปรปรวน ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา เขตพระโขนง กรุงเทพมหานคร จำนวน 1,066 คนที่ได้มาด้วยวิธีการสุ่มอย่างง่าย เครื่องมือที่ใช้ คือ แบบทดสอบความเข้าใจในการอ่าน จำนวน 50 ข้อ ชนิด 5 ตัวเลือก ที่ผู้วิจัยสร้างขึ้นเอง

ผลการวิจัยพบว่า การวิเคราะห์ความลำเอียงด้วยวิธีแปลงค่าความยาก กับวิธีไคสแควร์ และวิธีแปลงค่าความยาก กับวิธีวิเคราะห์ความแปรปรวนมีจำนวนข้อที่มีความลำเอียงแตกต่างกันมากอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีไคสแควร์ กับวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อที่มีความลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าความเชื่อมั่นของแบบทดสอบหลัง

คัดข้อที่ลำเอียงออกแล้วที่คำนวณด้วยสูตร KR-20 และแบบแบ่งครึ่งฉบับที่วิเคราะห์ด้วยวิธีทั้ง 3 พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

อังคณา สายยศ (2539:บทคัดย่อ) ได้ทำการวิจัยเรื่อง “การศึกษาผลการวิเคราะห์ความลำเอียงของแบบทดสอบความเข้าใจในการอ่านตามลักษณะการใช้ภาษาพูดในครอบครัว” โดยกลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัด สุรินทร์ ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น อำเภอเป็นชั้นและโรงเรียนเป็นหน่วยการสุ่ม ได้จำนวนโรงเรียนทั้งสิ้น 29 โรงเรียน ได้จำนวนนักเรียน 860 คน เครื่องมือที่ใช้ เป็นแบบทดสอบความเข้าใจในการอ่าน จำนวน 40 ข้อ ทำการวิเคราะห์ความลำเอียงด้วยวิธีแมนเทล-แฮนเดลเซล

ผลการวิจัยพบว่า จำนวนข้อสอบที่ลำเอียงของแบบทดสอบความเข้าใจในการอ่านระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดเป็น ภาษาลาว ภาษาเขมร และภาษาส่วย แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับความเชื่อมั่นของแบบทดสอบหลังคัดข้อลำเอียงออกแล้วพบว่า ค่าความเชื่อมั่นระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดเป็น ภาษาลาว ภาษาเขมร และภาษาส่วยแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนคะแนนที่ได้จากการสอบด้วยแบบทดสอบที่คัดข้อลำเอียงออกแล้วมีคะแนนระหว่างกลุ่มผู้สอบที่พูดภาษาลาวกับภาษาส่วย ภาษาเขมรกับภาษาส่วยมีค่าความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนภาษาลาวกับภาษาเขมรมีคะแนนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

คมศักดิ์ ชื่นชม (2539:78-82) ได้ทำการวิจัยเรื่อง “การศึกษาผลการวิเคราะห์ความลำเอียงที่ใช้วิธีต่างกันของแบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์” กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2538 ของโรงเรียนในจังหวัด ยะลา ซึ่งได้มาโดยวิธีการสุ่มแบบแบ่งชั้น โดยใช้ภาษาที่นักเรียนใช้ในชีวิตประจำวันเป็นชั้น จำนวน 1,081 คน แบ่งออกเป็น 2 กลุ่ม กลุ่มที่ 1 ใช้ในการเปรียบเทียบผลการวิเคราะห์ความลำเอียงจำนวน 800 คน กลุ่มที่ 2 ใช้เปรียบเทียบความเชื่อมั่นของแบบทดสอบหลังคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จำนวน 281 คน เครื่องมือที่ใช้ คือ แบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์ จำนวน 30 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง ระหว่างกลุ่มผู้สอบที่ใช้ภาษาในชีวิตประจำวันในครอบครัวต่างกัน คือ ภาษาไทยและไม่ใช้ภาษาไทย และนับถือศาสนาต่างกัน คือ ศาสนาพุทธ ศาสนาอิสลาม ด้วยวิธีวิเคราะห์ 3 วิธี คือ วิธีค่าอำนาจจำแนก วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอยโดยทำการเปรียบเทียบจำนวนข้อสอบที่ลำเอียง และศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบ

ผลการวิจัยพบว่า การวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอย มีจำนวนข้อที่มีความลำเอียงน้อยและไม่แตกต่างกันที่ระดับนัยสำคัญทางสถิติ .05 ค่าความเชื่อมั่นของแบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์ หลังคัดเลือกข้อที่มีความลำเอียงออก จากการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีค่าอำนาจจำแนกของข้อสอบ วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอยมีค่าไม่แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .05

เรวดี อินทสระ (2539:248-252) ได้ทำการวิจัยเรื่อง “ผลการตรวจสอบความลำเอียงข้อสอบต่อการศึกษาคำความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน” โดยศึกษาเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบ 3 วิธี คือ วิธีแปลงค่าความยาก วิธีโค้งลักษณะข้อสอบ และวิธีแมนเทล-แฮนเดลเซล ของข้อคัดเลือกเข้าศึกษาในชั้นปีที่ 1 ประเภทรับตรง ปีการศึกษา 2538 ของมหาวิทยาลัย สงขลานครินทร์ ในวิชาภาษาไทย ก วิชาสังคมศึกษา ก วิชาภาษาอังกฤษ กข วิชาละ 8,127 คน (ชาย 2,722 คน หญิง 5,405 คน) วิชาภาษาไทย กข วิชาสังคมศึกษา กข และวิชาภาษาอังกฤษ กขค วิชาละ 5,415 คน (ชาย 1,454 คน หญิง 3,961 คน) ความเที่ยงตรงเชิงพยากรณ์ศึกษาจากคะแนนสอบคัดเลือกกับระดับผลการเรียนภาคเรียนที่ 1 ปีการศึกษา 2538 ของนักศึกษาที่ได้รับการคัดเลือกจากประเภทรับตรง

สายวิทยาศาสตร์ 763 คน และสายศิลปศาสตร์ 281 คน และสาเหตุความลำเอียงของข้อสอบจากการระบุสาเหตุของนักวัดผลการศึกษา หรืออาจารย์ผู้สอนจำนวน 50 คน และนิสิตนักศึกษาที่เรียนในสาขาวิชานั้นๆ วิชาละ 50 คน

ผลการวิจัยพบว่า วิธีตรวจสอบความลำเอียงทั้ง 3 วิธี ตัดสินจำนวนข้อสอบที่ลำเอียงแตกต่างกันในวิชาภาษาไทย ก ฉบับที่ 2 และวิชาสังคมศึกษา ก ฉบับที่ 1 ที่ระดับนัยสำคัญทางสถิติที่ .05 นอกนั้นต่างกันที่ระดับนัยสำคัญทางสถิติที่ .01 โดยใช้วิธีทฤษฎีการตอบข้อคำถาม ตัดสินจำนวนข้อสอบที่ลำเอียงได้มากที่สุด ความสัมพันธ์ของลำดับที่ของการสอบไม่ว่าจะคิดคะแนนมาตรฐานที่-ปกติ หรือ คิดคะแนนน้ำหนักความสามารถ และใช้ข้อสอบจำนวนทั้งหมดหรือใช้เฉพาะข้อสอบที่ปราศจากความลำเอียงต่างมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ประสิทธิภาพในการทำนายผลสัมฤทธิ์ทางการเรียนในสายวิทยาศาสตร์ การคิดคะแนนน้ำหนักความสามารถทั้งใช้ข้อสอบจำนวนทั้งหมด และใช้เฉพาะข้อสอบที่ปราศจากความลำเอียงมีประสิทธิภาพในการทำนายสูงกว่าการคิดคะแนนมาตรฐานที่-ปกติ ส่วนสายศิลปศาสตร์ การคิดคะแนนมาตรฐานที่-ปกติทั้งใช้ข้อสอบจำนวนทั้งหมดและใช้เฉพาะข้อสอบที่ปราศจากความลำเอียง มีประสิทธิภาพในการทำนายสูงกว่า การคิดคะแนนน้ำหนักความสามารถ และสาเหตุของความลำเอียงของข้อสอบต่อเพศทั้งชายและหญิง เกิดจากข้อสอบเป็นข้อคำถามที่ผู้สอบได้รับการฝึกฝนเฉพาะจะมีโอกาสตอบถูกมากกว่า เป็นเรื่องราวที่กลุ่มนั้นๆสนใจ และเป็นข้อสอบที่ถามความจำ

จิตสุตา ธารพร (2539:129-134) ได้ทำการวิจัยเรื่อง "การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ และแบบทดสอบวัดความเข้าใจทางภาษา" ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2538 โรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดกระบี่ จำนวน 2,689 คน ได้มาโดยวิธีการสุ่มอย่างง่าย เครื่องมือที่ใช้มี 2 ฉบับ ฉบับที่ 1 คือ แบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ เรื่อง สมการและอสมการ อัตราส่วนร้อยละ ปริมาตรและพื้นที่ผิว แบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 40 ข้อ ฉบับที่ 2 คือ แบบทดสอบความเข้าใจทางภาษา วิชา ภาษาไทย แบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 40 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง ส่วนวิธีวิเคราะห์มี 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีแมนเทิล-แฮนด์เชล

ผลการวิจัยพบว่า วิธีวิเคราะห์ของแมนเทิล-แฮนด์เชล พบจำนวนข้อสอบที่มีความลำเอียงมากที่สุดจากแบบทดสอบทั้ง 2 ฉบับ และวิธีแปลงค่าความยากของข้อสอบพบจำนวนข้อสอบที่มีความลำเอียงน้อยที่สุด จำนวนข้อสอบที่มีความลำเอียงในแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์นั้น พบว่า การวิเคราะห์ด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธีแมนเทิล-แฮนด์เชล มีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีวิเคราะห์โดยโค้งลักษณะข้อสอบกับวิธีแมนเทิล-แฮนด์เชล มีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และจำนวนข้อสอบที่มีความลำเอียงในแบบทดสอบความเข้าใจทางภาษา พบว่า การวิเคราะห์ด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธีโค้งลักษณะข้อสอบ และการวิเคราะห์ความลำเอียงด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธีแมนเทิลแฮนด์เชลมีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ส่วนการวิเคราะห์ความลำเอียงด้วยวิธีโค้งลักษณะข้อสอบกับวิธีแมนเทิล-แฮนด์เชล มีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ส่วนค่าความเชื่อมั่นของแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ ที่คำนวณโดยสูตรแบบ KR-20 หลังจากคัดข้อสอบที่มีความลำเอียงออกด้วย วิธีแปลงค่าความยากของข้อสอบกับวิธีโค้งลักษณะข้อสอบ และวิธีแปลงค่าความยากของข้อสอบ กับ วิธีของแมนเทิล-แฮนด์เชล มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่

ระดับ .05 ส่วนโค้งลักษณะข้อสอบกับวิธีแบบแมนเทิล-แฮนด์เชล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นที่คำนวณโดยใช้สูตรแบบแบ่งครึ่งฉบับ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว พบว่าทั้ง 3 วิธี มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และสำหรับค่าความเชื่อมั่นของแบบทดสอบความเข้าใจทางภาษาที่คำนวณโดยใช้สูตรแบบ KR-20 และแบบแบ่งครึ่งฉบับ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว พบว่า ทั้ง 3 วิธี มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

พรรณี จินตมาศ (2540:90-92) ได้ทำการวิจัยเรื่อง “การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบ โดยใช้ขนาดกลุ่มผู้สอบและวิธีวิเคราะห์ต่างกัน” ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง จำนวน 2,200 คน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น มีขนาดของโรงเรียนเป็นชั้นและโรงเรียนเป็นหน่วยการสุ่ม เครื่องมือที่ใช้ คือ แบบทดสอบคณิตศาสตร์โจทย์ปัญหา เรื่องสมการ อัตราส่วน และร้อยละ แบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 40 ข้อ 1 ฉบับ ที่ผู้วิจัยสร้างขึ้นเอง โดยวิธีวิเคราะห์ความลำเอียง 3 วิธี คือ วิธีแปลงค่าความยาก วิธีแมนเทิล-แฮนด์เชล และวิธี ชิปเทสต์

ผลการวิจัยพบว่า เมื่อวิเคราะห์จากกลุ่มผู้สอบขนาด 500 คน วิธีชิปเทสต์พบข้อสอบที่มีความลำเอียงมากที่สุด และวิธีแปลงค่าความยากพบข้อสอบที่มีความลำเอียงน้อยที่สุด โดยจำนวนข้อสอบที่ลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทุกวิธี เมื่อวิเคราะห์กลุ่มผู้สอบขนาด 1,000 คน วิธีแมนเทิล-แฮนด์เชล พบข้อสอบที่มีความลำเอียงมากที่สุด วิธีแปลงค่าความยากไม่พบข้อสอบที่ลำเอียง โดยจำนวนข้อสอบที่ลำเอียงจากการวิเคราะห์ด้วยวิธีแปลงค่าความยากกับวิธีแมนเทิลแฮนด์เชล และวิธีแปลงค่าความยากกับวิธีชิปเทสต์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกจากนั้นมีค่าความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับจำนวนข้อสอบที่ลำเอียงจากการวิเคราะห์ด้วยวิธีแปลงค่าความยากระหว่างกลุ่มผู้สอบขนาด 500 คน และกลุ่มผู้สอบขนาด 1,000 คน จะมีจำนวนข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบหลังคัดเลือกข้อสอบที่มีความลำเอียงออกเมื่อวิเคราะห์จากกลุ่มผู้สอบขนาด 500 คน ด้วยวิธีแปลงค่าความยากกับวิธีชิปเทสต์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และเมื่อวิเคราะห์จากกลุ่มผู้สอบขนาด 1,000 คน ด้วยวิธีวิเคราะห์ความลำเอียงทั้ง 3 วิธี มีค่าความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

อารี วัชรโสติกุล (2543:55-59) ได้ทำการวิจัยเรื่อง “การเปรียบเทียบผลการทำหน้าที่ต่างกันของข้อสอบโดยใช้รูปแบบและวิธีการต่างกัน” โดยศึกษาเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้รูปแบบต่างกัน คือ รูปแบบคะแนนรวมทั้งฉบับ แยกตามเนื้อหา และแยกตามระดับพฤติกรรม ด้วยวิธีการตรวจสอบต่างกัน คือ วิธีชิปเทสต์ และวิธีถดถอยโลจิสติก แล้วทำการคัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบทดสอบ เพื่อเปรียบเทียบ ค่าความเชื่อมั่น กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา กรุงเทพมหานคร ส่วนกลาง กลุ่ม 4 จำนวน 994 คน เป็นนักเรียนชาย 480 คน นักเรียนหญิง 514 คน เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 50 ข้อ

ผลการวิจัยพบว่า จำนวนข้อสอบที่ทำหน้าที่ต่างกันโดยใช้รูปแบบต่างกันแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จำนวนข้อสอบที่ทำหน้าที่ต่างกันโดยใช้วิธีการตรวจสอบต่างกันแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในรูปแบบรวมทั้งฉบับ และแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติในรูปแบบแยกตามเนื้อหา และรูปแบบแยกตามระดับพฤติกรรม ส่วนค่าความเชื่อมั่นของแบบทดสอบหลังคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกโดยใช้รูปแบบการตรวจสอบต่างกัน เมื่อใช้วิธีตรวจสอบด้วยวิธีชิปเทสต์ พบว่า ค่าความ

เชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แล้วจึงทดสอบค่าความเชื่อมั่นเป็นรายคู่ พบว่า ค่าความเชื่อมั่นระหว่างรูปแบบแยกตามเนื้อหาและรูปแบบแยกตามระดับพฤติกรรม แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการตรวจสอบด้วยวิธีถดถอยโลจิสติก พบว่า ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และเมื่อตรวจสอบโดยใช้รูปแบบรวมทั้งฉบับ และ รูปแบบแยกตามเนื้อหาพบว่า ค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อตรวจสอบโดยใช้รูปแบบแยกตามระดับพฤติกรรม พบว่า ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

วิภา จำมัน (2544:49-51) ได้ทำการวิจัยเรื่อง "การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษาตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน" โดยศึกษาเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงและค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 1,041 คน ได้มาโดยการสุ่มแบบแบ่งชั้น

ผลการวิจัยพบว่า จำนวนข้อสอบที่มีความลำเอียง เมื่อตรวจสอบด้วยวิธีชิปเทสต์ ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนข้อสอบที่มีระดับความง่ายต่ำกับระดับความง่ายปานกลาง และข้อสอบที่มีระดับความง่ายปานกลางกับระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เซล พบว่า ข้อสอบที่มีระดับความง่ายต่ำ ปานกลาง สูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อตรวจสอบระหว่างวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนด์เซล พบว่า ข้อสอบที่มีระดับความง่ายต่ำ ปานกลาง สูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นของแบบทดสอบหลังคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ของข้อสอบที่มีระดับความง่ายต่ำ ปานกลาง สูง เมื่อตรวจสอบระหว่างวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนด์เซล พบว่า มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากเอกสารและงานวิจัยที่เกี่ยวข้องทั้งต่างประเทศ และในประเทศที่กล่าวมาข้างต้น พบว่าการศึกษาเปรียบเทียบวิธีการทำหน้าที่ต่างกันของข้อสอบมีหลายวิธี เช่น วิธีแปลงค่าความยาก วิธีค่าอำนาจจำแนก วิธีไคสแควร์ วิธีวิเคราะห์ความแปรปรวน วิธีโค้งลักษณะข้อสอบ วิธีแมนเทิล-แฮนด์เซล วิธีชิปเทสต์ เป็นต้น จากวิธีการดังกล่าว บางวิธีให้ผลการวิเคราะห์ที่ยุ่งยาก เสียค่าใช้จ่ายสูง บางวิธีใช้กลุ่มตัวอย่างจำนวนไม่มาก การคำนวณและการแปลผลง่าย มีความถูกต้อง ซึ่งมีผู้นิยมนำมาใช้ ได้แก่ วิธีแมนเทิล-แฮนด์เซล วิธีวิเคราะห์ความแปรปรวน สำหรับตัวแปรที่นำมาใช้ในการศึกษาการทำหน้าที่ต่างกันของข้อสอบที่นิยมทำการศึกษา มี ดังนี้ ภาษา เพศ วัฒนธรรม ศาสนา ภูมิศาสตร์ เป็นต้น แต่การศึกษาการทำหน้าที่ต่างกันของ ข้อสอบในรูปแบบของความคุ้นเคย ความสนใจ ความพอใจ ที่มีต่อแบบทดสอบวัดความถนัดทางการเรียนนั้นยังไม่ปรากฏว่ามีผู้ศึกษาการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีวิเคราะห์แบบแมนเทิล-แฮนด์เซล และวิธีวิเคราะห์ความแปรปรวน ดังนั้น ผู้วิจัยมีความสนใจที่จะศึกษาการทำหน้าที่ต่างกันของข้อสอบโดยมี รูปแบบ และวิธีดังกล่าว

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ดำเนินการตามลำดับดังนี้

1. การกำหนดประชากร และกลุ่มตัวอย่าง
2. การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า
3. การเก็บรวบรวมข้อมูล
4. การจัดกระทำข้อมูลและการวิเคราะห์ข้อมูล

1. การกำหนดประชากร และกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ จำนวน โรงเรียน 53 โรงเรียน จำนวน 103 ห้องเรียน รวมนักเรียนทั้งสิ้น 2,855 คน จำแนกเป็นนักเรียนชาย 1,491 คน นักเรียนหญิง 1,364 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน (Two-Stage Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) โรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) จำนวนโรงเรียน 9 โรงเรียน จำนวน 22 ห้องเรียน รวมนักเรียนทั้งสิ้น 585 คน จำแนกเป็นนักเรียนชาย 310 คน นักเรียนหญิง 275 คน โดยกำหนดกลุ่มตัวอย่างทั้งหมดด้วยความเชื่อมั่น 95% ($\alpha = .05$) ซึ่งเทียบจากตารางขนาดกลุ่มตัวอย่างของ ยามาเน่ (ล้วน สายยศ และอังคณา สายยศ. 2538 : 298 อ้างอิงจาก Yamane, 1967) ปรากฏว่าต้องใช้ขนาดกลุ่มตัวอย่างอย่างน้อย 345 คน ในการศึกษาครั้งนี้ มีลำดับขั้นตอนในการสุ่มดังนี้

ขั้นที่ 1 สุ่มโรงเรียนจากแต่ละขนาดของโรงเรียนด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดของโรงเรียนเป็นชั้น (Strata) คือ โรงเรียนขนาดใหญ่ (พิจารณาจากจำนวนนักเรียนทั้งหมดมากกว่า 800 คน) โรงเรียนขนาดกลาง (พิจารณาจำนวนนักเรียนทั้งหมดมากกว่า 401 คน แต่น้อยกว่า 800 คน) โรงเรียนขนาดเล็ก (พิจารณาจากจำนวนนักเรียนทั้งหมดน้อยกว่า 400 คน) และมีขนาดโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้โรงเรียนขนาดใหญ่ 2 โรงเรียน โรงเรียนขนาดกลาง 4 โรงเรียน และโรงเรียนขนาดเล็ก 3 โรงเรียน

ขั้นที่ 2 สุ่มห้องเรียนจากโรงเรียนที่สุ่มได้ในแต่ละขนาด โดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) มีห้องเรียนเป็นหน่วยการสุ่ม ได้จำนวนห้องเรียน 20 ห้องเรียน จำนวนนักเรียน 585 คน จำแนกเป็นนักเรียนชาย 310 คน นักเรียนหญิง 275 คน ดังตาราง 3

ตาราง 3 แสดงจำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามขนาดโรงเรียน

ขนาด	โรงเรียน	จำนวนนักเรียน(ห้องเรียน)		
		ชาย	หญิง	รวม(ห้อง)
ใหญ่	โรงเรียนวัดจันทร์สโมสร	30	38	68(2)
	โรงเรียนวัดไผ่ตัน	80	59	139(5)
กลาง	โรงเรียนวัดมกุฏกษัตริยาราม	44	34	78(2)
	โรงเรียนปทุมวัน	40	32	72(2)
	โรงเรียนวัดประชาศรีธรรม	26	31	57(2)
	โรงเรียนวัดพระยาอภัย	24	33	57(2)
เล็ก	โรงเรียนวัดอินทวิหาร	19	15	34(2)
	โรงเรียนวัดหัวลำโพง	22	15	37(1)
	โรงเรียนวัดเทวราชกุญชร	25	18	43(2)
รวม		310	275	585(20)

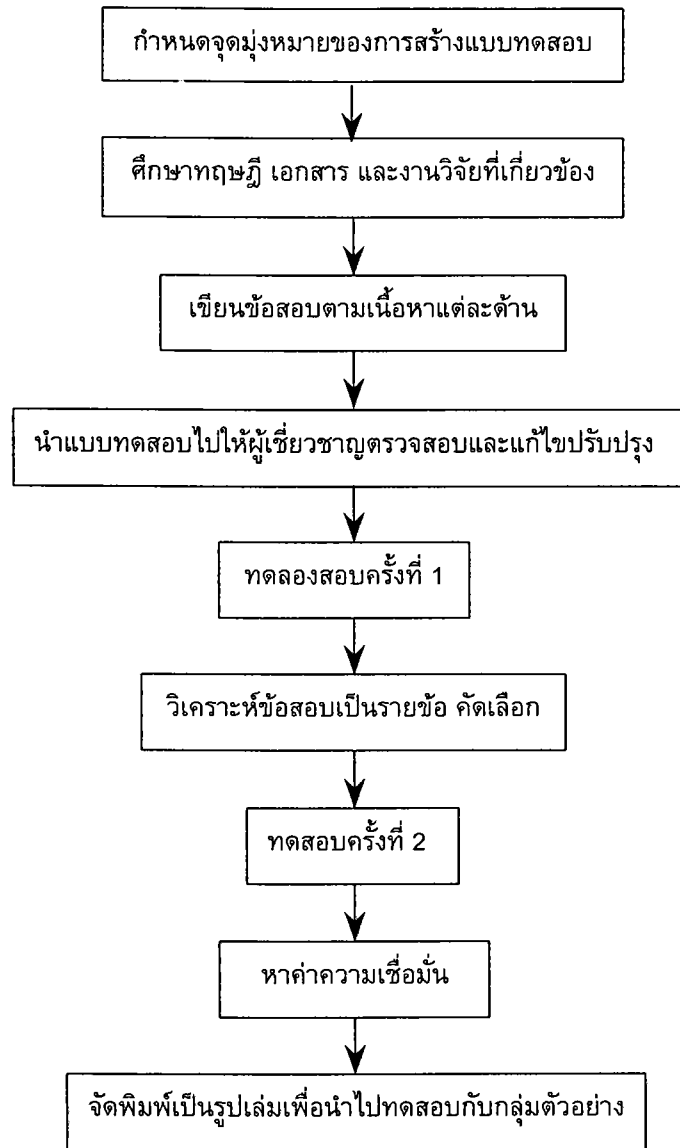
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดความถนัดทางการเรียนที่สร้างโดย
ยึดหลักทฤษฎีหลายองค์ประกอบ ตามแนวของเทอร์สโตน (Thurstone) ดังนี้

1. แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล (R) (ข้อที่ 1 ถึง 30)
2. แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา (V) (ข้อที่ 31 ถึง 60)
3. แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข (N) (ข้อที่ 61 ถึง 90)

เป็นข้อสอบปรนัยชนิด 5 ตัวเลือก ที่ผู้วิจัยสร้างขึ้นเอง

วิธีดำเนินการสร้างเครื่องมือ



ภาพประกอบ 5 แสดงลำดับขั้นตอนในการสร้างเครื่องมือ

ผู้วิจัยได้ดำเนินการสร้างเครื่องมือสำหรับใช้ในงานวิจัยตามภาพประกอบข้างต้น ดังมีรายละเอียดดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อสร้างแบบทดสอบวัดความถนัดทางการเรียน ด้านตัวเลข ด้านเหตุผล และด้านภาษาที่ใช้ในศึกษาเปรียบเทียบผลการวิเคราะห์การทำหน้าที่ต่างกันของแบบทดสอบ

2. ศึกษา ทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้องกับทฤษฎีหลายองค์ประกอบของเทอร์สโตน (Thurstone) เพื่อเป็นแนวทางในการสร้างแบบทดสอบในงานวิจัย

3. เขียนข้อสอบวัดความถนัดทางการเรียน ในลักษณะข้อสอบปรนัย ชนิด 5 ตัวเลือก ดังนี้

3.1 แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จำนวน 45 ข้อ

3.2 แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จำนวน 45 ข้อ

3.3 แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จำนวน 45 ข้อ

4. เมื่อสร้างแบบทดสอบเสร็จแล้ว ได้นำไปให้ผู้เชี่ยวชาญ ตรวจสอบ เพื่อนำมาปรับปรุงแก้ไขความถูกต้องด้านเนื้อหาและภาษาที่ใช้ โดยพิจารณาจากค่า IOC ของข้อสอบแต่ละข้อที่คัดเลือก ให้มีค่ามากกว่าหรือเท่ากับ 0.5 ซึ่งปรากฏดังนี้

4.1 แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จำนวน 45 ข้อ ผ่านการคัดเลือกจำนวน 32 ข้อ โดยมีค่า IOC ระหว่าง 0.60 ถึง 1.00 และมีการปรับปรุงข้อสอบจำนวน 13 ข้อ ที่มีค่า IOC ต่ำกว่า 0.50 รวมได้แบบทดสอบที่ปรับปรุงแล้วจำนวน 45 ข้อ

4.2 แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จำนวน 45 ข้อ ผ่านการคัดเลือกจำนวน 27 ข้อ โดยมีค่า IOC ระหว่าง 0.60 ถึง 0.80 และมีการปรับปรุงข้อสอบจำนวน 18 ข้อ รวมได้แบบทดสอบที่ปรับปรุงแล้วจำนวน 45 ข้อ

4.3 แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จำนวน 45 ข้อ ผ่านการคัดเลือกจำนวน 39 ข้อ โดยมีค่า IOC ระหว่าง 0.60 ถึง 1.00 และมีการปรับปรุงข้อสอบจำนวน 6 ข้อ รวมได้แบบทดสอบที่ปรับปรุงแล้วจำนวน 45 ข้อ

5. นำแบบทดสอบวัดความถนัดทางการเรียน ที่ปรับปรุงแก้ไขแล้วไปทดสอบครั้งที่ 1 กับนักเรียนชั้นประถมศึกษาปีที่ 6 จากโรงเรียนวัดบางปะกอก จำนวน 106 คน และโรงเรียนราษฎร์บูรณะ จำนวน 135 คน รวมทั้งหมด จำนวน 241 คน แล้วนำมาตรวจให้คะแนน โดยข้อที่ทำถูกต้อง 1 คะแนน ข้อที่ทำผิดได้ 0 คะแนน จากนั้นจึงนำมาวิเคราะห์รายข้อ โดยนำคะแนนมาเรียงลำดับแยกเป็นกลุ่มสูง กลุ่มต่ำ โดยใช้เทคนิค 27 % แล้วจึงหาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เพื่อคัดเลือกให้ได้ข้อสอบที่มีค่าความยากง่ายตั้งแต่ 0.20 ถึง 0.80 และมีค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป ปรากฏว่าได้ข้อสอบที่คัดเลือกแล้วด้านละ 30 ข้อ รวมจำนวนข้อสอบทั้งหมด 90 ข้อ มีค่าความยากง่ายระหว่าง 0.216 ถึง 0.805 และมีค่าอำนาจจำแนกระหว่าง 0.204 ถึง 0.587 ดังรายละเอียดในตาราง 11 ภาคผนวก ข ดังนี้

5.1 แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จำนวน 30 ข้อ มีค่าความยากง่ายระหว่าง 0.216 ถึง 0.784 มีค่าอำนาจจำแนกระหว่าง 0.204 ถึง 0.532

5.2 แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จำนวน 30 ข้อ มีค่าความยากง่ายระหว่าง 0.248 ถึง 0.805 มีค่าอำนาจจำแนกระหว่าง 0.204 ถึง 0.558

5.3 แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จำนวน 30 ข้อ มีค่าความยากง่ายระหว่าง 0.249 ถึง 0.743 มีค่าอำนาจจำแนกระหว่าง 0.245 ถึง 0.584

6. นำแบบทดสอบวัดความถนัดทางการเรียนที่คัดเลือก ไปทดลองสอบครั้งที่ 2 กับนักเรียนชั้นประถมศึกษาปีที่ 6 จากโรงเรียนวัดทุ่งครุ จำนวน 104 คน และโรงเรียนประชาภิบาล จำนวน 91 คน รวมทั้งหมดจำนวน 195 คน แล้วตรวจให้คะแนนเช่นเดียวกับข้อ 5 เพื่อนำมาหาค่าความเชื่อมั่นของแบบทดสอบแต่ละด้าน โดยวิธีหาค่าความสอดคล้องภายในของแบบทดสอบจากสูตร KR-20 ตามวิธีของคูเดอร์-ริชาร์ดสัน (บุญเชิด ภิญโญนนตพงษ์.2545:129) ซึ่งปรากฏว่าได้ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา และด้านตัวเลข เท่ากับ 0.748 , 0.699 และ 0.784 ตามลำดับ ซึ่งมีค่าความเชื่อมั่นสูงพอที่ใช้ในการวิจัยครั้งนี้

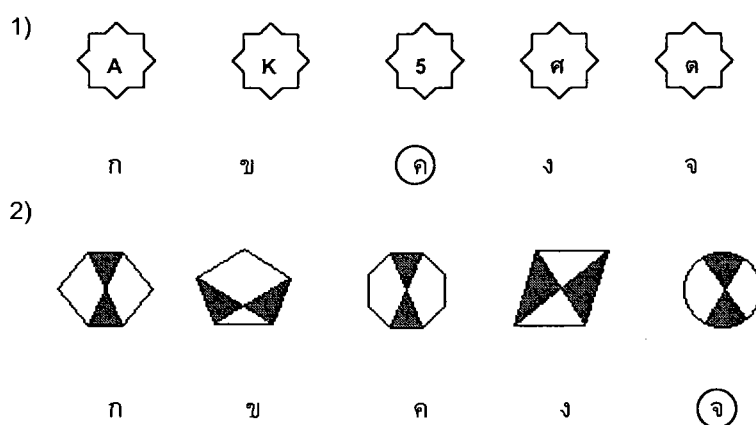
7. จากนั้นจึงนำแบบทดสอบที่ได้จากข้อ 6 มาจัดพิมพ์เป็นรูปเล่ม เพื่อนำไปทดสอบกับกลุ่มตัวอย่าง

ลักษณะเนื้อหาของแบบทดสอบ

แบบทดสอบที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบวัดความถนัดทางการเรียนมีลักษณะ และรายละเอียด ดังนี้

1. แบบทดสอบความถนัดทางการเรียนด้านเหตุผล (R) เป็นแบบทดสอบวัดความสามารถในการคิดหาเหตุผลเกี่ยวกับการหาค่า รูปภาพที่ไม่เข้าพวก และการสรุปความ เป็นข้อสอบปรนัยชนิด 5 ตัวเลือก จำนวน 30 ข้อ ที่มีคำตอบถูกเพียงคำตอบเดียว ดังนี้

1.1 ชนิดจัดประเภทรูปภาพ ข้อสอบแต่ละข้อจะกำหนดรูป มาให้แล้วให้พิจารณารูปที่แตกต่างกัน ตัวอย่างเช่น



1.2 ชนิดจัดประเภทภาษา ข้อสอบแต่ละข้อจะกำหนดคำมาให้แล้วให้พิจารณาว่าคำในข้อใดไม่เข้าพวก ตัวอย่างเช่น

1) ข้อใดไม่เข้าพวก

- ก. รถไฟ
- ข. รถยนต์
- ค. เรือยนต์
- ง. เครื่องบิน
- จ. รถจักรยาน

2) ข้อใดไม่เข้าพวก

- ก. เทนนิส
- ข. ฟุตบอล
- ค. บาสเกตบอล
- ง. วอลเลย์บอล
- จ. แบดมินตัน

1.3 ชนิดสรุปความ ข้อสอบแต่ละข้อ จะมีข้อความมาให้ แล้วให้สรุปความจากข้อความนั้นๆ ตัวอย่างเช่น

1) เจยืนอยู่หน้าใจ จิมยืนอยู่หลังแจนแต่อยู่หน้าใจ ใครยืนอยู่หลังสุด?

- ก. เจ
- ข. ใจ
- ค. จิม
- ง. แจน
- จ. ยังสรุปไม่ได้

2) ตึกวิ่งเร็วกว่าตึก แต่ช้ากว่าต็อก ต็อกวิ่งเร็วกว่าตึก แต่ช้ากว่าต๋อย ใครวิ่งเร็วที่สุด?

- ก. ตึก
- ข. ตึก
- ค. ต็อก
- ง. ตึก
- จ. ต๋อย

2. แบบทดสอบความถนัดทางการเรียนด้านภาษา (V) เป็นแบบทดสอบวัดความสามารถทางภาษา เกี่ยวกับการหาความหมาย คำตรงข้าม ศัพท์สัมพันธ์ และความเข้าใจภาษา เป็นข้อสอบปรนัยชนิด 5 ตัวเลือก จำนวน 30 ข้อ ที่มีคำตอบถูกเพียงคำตอบเดียว ดังนี้

2.1 ชนิดคำตรงข้าม ข้อสอบแต่ละข้อจะกำหนดคำมาให้แล้วให้นักเรียนหาคำตอบที่มีความหมายตรงข้ามกับคำที่กำหนด หรือคำที่ขีดเส้นใต้ในประโยค ตัวอย่างเช่น

1) น้องหนูเป็นเด็กอ่อนโยน

- ก. ยึดมั่น
- ข. กระด้าง
- ค. หยาดคาย
- ง. หนักแน่น
- จ. เข้มแข็ง

2) คนมาชุมนุมกันหนาแน่น

- ก. เบาบาง
- ข. เจือจาง
- ค. เล็กน้อย
- ง. หละหลวม
- จ. ผอมบาง

2.2 ศัพท์สัมพันธ์ ข้อสอบแต่ละข้อจะกำหนดคำมาให้แล้วให้นักเรียนหาคำตอบที่มีความหมายสัมพันธ์กับคำที่กำหนดให้ ตัวอย่างเช่น

1) ทหาร

- ก. ลูก
- ข. ขโมย
- ค. ข้าศึก
- ง. ผู้ร้าย
- จ. ห้องขัง

2) ลอย

- ก. นุ่ม
- ข. เบา
- ค. บาง
- ง. อ่อน
- จ. ละออง

2.3 ความเข้าใจภาษา ข้อสอบแต่ละข้อจะกำหนดข้อความ ประโยค คำพังเพยมาให้แล้วให้นักเรียนอ่านแล้วเลือกคำที่ถูกต้องเหมาะสม ดังนี้

..ฝ่ายขุนช้างหมางจิตให้คิดแค้น ลูกขุนแผนมันคงไม่สงสัย
เมื่อกระนั้นเหมือนกูครั้นดูไป ก็กลับไพร่เหมือนพ่ออำยทรพี...

1) ข้อความข้างบนนี้ควรเป็นข้อความประเภทใด

- ก. นิဂ
- ข. ด่า
- ค. ชม
- ง. พุด
- จ. ป่น

2) ข้อความใดจะเหมาะสมกับคำกลอนมากที่สุด

- ก. เสียชีพอยาเสียสัตย์
- ข. ที่ได้มีรักที่นั่นมีทุกข์
- ค. รักอื่นใดจักเสมอรักตนเอง
- ง. อดเยี่ยงอย่างเสือสงวนศักดิ์
- จ. ☒ รวมกันเราอยู่แยกกันอยู่เราตาย

3. แบบทดสอบความถนัดทางการเรียนด้านตัวเลข (N) เป็นแบบทดสอบวัดความสามารถในการคิดคำนวณเกี่ยวกับอนุกรมตัวเลข และคณิตศาสตร์เหตุผล เป็นข้อสอบปรนัยชนิด 5 ตัวเลือก จำนวน 30 ข้อ ที่มีคำตอบถูกเพียงคำตอบเดียว ดังนี้

3.1 ตัวเลขอนุกรมธรรมดา ข้อสอบแต่ละข้อจะเป็นอนุกรมในแนวเดียว ที่สามารถได้ในระบบเดียว และระบบซ้อน ดังตัวอย่างเช่น

1) 8 4 12 15 5 20 24

- ก. 2
- ข. 4
- ค. 5
- จ. ☒ 6
- ง. 8

2) 23 30 24 31

- ก. ☒ 25
- ข. 26
- ค. 32
- ง. 33
- จ. 35

3.2 คณิตศาสตร์เหตุผล ข้อสอบแต่ละข้อจะกำหนดโจทย์มาให้ แล้วให้พิจารณาแก้โจทย์ปัญหาคณิตศาสตร์ ดังตัวอย่างเช่น

1) B หาร A มีค่าอยู่ระหว่าง 3 กับ 6 ค่าของ A และ B ควรเป็นข้อใด

- ก. $A = 10$, $B = 5$
- ข. $A = 24$, $B = 4$
- ค. ☒ $A = 25$, $B = 5$
- ง. $A = 42$, $B = 6$
- จ. ไม่มีคำตอบข้อใดถูก

2) ชื่อของมาราคา A บาท ขายไปราคา B บาท ได้กำไรเท่าไร

- ก. A
ข. B
ค. A+B
ง. A-B
จ. B-A

คำชี้แจงในการตอบข้อสอบ

1. แบบทดสอบทั้งหมดจำนวน 3 ฉบับ ฉบับละ 30 ข้อ
2. ข้อสอบแต่ละข้อเป็นข้อสอบปรนัย 5 ตัวเลือก ก. , ข. , ค. , ง. และ จ. ให้นักเรียนพิจารณาว่าตัวเลือกใดเป็นตัวเลือกถูกต้องที่สุดให้กากบาทลงในกระดาษคำตอบดังตัวอย่าง
3. แล้วให้นักเรียนพิจารณาว่า
 - ข้อสอบข้อนี้นักเรียนได้เคยทำ หรือเคยได้รับการฝึกทำโจทย์ลักษณะเช่นนี้มาก่อน ให้นักเรียนกากบาทลงในช่องคุ้นเคย หากว่าไม่เคยให้กากบาทลงในช่องไม่คุ้นเคย
 - ข้อสอบข้อนี้นักเรียนมีความรู้สึกชอบ อยากที่จะทำ ให้นักเรียนกากบาทลงในช่องสนใจ หากว่าไม่ใช่ให้กากบาทลงในช่องไม่สนใจ
 - ข้อสอบข้อนี้นักเรียนมีความพร้อม ยินดี ที่จะทำ ให้นักเรียนกากบาทลงในช่องพอใจ หากว่าไม่ใช่ให้กากบาทลงในช่องไม่พอใจ

ตัวอย่าง

1) เจยืนอยู่หน้าใจ จิมยืนอยู่หลังแจนแต่อยู่หน้าใจ ใครยืนอยู่หลังสุด?

- ก. เจ
ข. ใจ
ค. จิม
ง. แจน
จ. ยังสรุปไม่ได้

ลักษณะของกระดาษคำตอบ

ข้อ	คำตอบ					ความคุ้นเคย		ความสนใจ		ความพอใจ	
						คุ้นเคย	ไม่คุ้นเคย	สนใจ	ไม่สนใจ	พอใจ	ไม่พอใจ
1	ก	☒	ค	ง	จ		☒	☒			☒
2	ก	ข	ค	ง	จ						
3	ก	ข	ค	ง	จ						
4	ก	ข	ค	ง	จ						
5	ก	ข	ค	ง	จ						

1. กำหนดเวลาที่จะทำการทดสอบ โดยติดต่อขออนุญาตผู้บริหารโรงเรียนที่เลือกเป็นกลุ่มตัวอย่าง
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่จะทำการสอบในแต่ละครั้ง วางแผนดำเนินการสอบ
3. ชี้แจงทำความเข้าใจกับผู้ประสานในแต่ละโรงเรียน ในเรื่องลักษณะของแบบทดสอบ และวิธีดำเนินการสอบ
4. นำแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา และด้านตัวเลข ไปทดสอบกับกลุ่มตัวอย่างจำนวน 585 คน เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน และทำการทดสอบสมมติฐานการศึกษาค้นคว้า

4. การจัดกระทำข้อมูลและการวิเคราะห์ข้อมูล

ผู้วิจัยทำการวิเคราะห์ข้อมูล คำนวณโดยใช้โปรแกรม SIBTEST โปรแกรม SPSS for WINDOWS (Statistic Package for Social Sciences for WINDOWS) และโปรแกรมคอมพิวเตอร์ที่ใช้ในการวิเคราะห์ ข้อสอบ (CTIA) โดยมีขั้นตอนในการวิเคราะห์ข้อมูล และสถิติที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

ขั้นตอนในการวิเคราะห์ข้อมูล

ผู้วิจัยได้ดำเนินการตามขั้นตอนต่อไปนี้

1. วิเคราะห์ข้อมูลในขั้นตอนการทดลองเครื่องมือ ได้ดำเนินการ ดังนี้
 - 1.1 ตรวจสอบให้คะแนนตามเกณฑ์ที่ตั้งไว้
 - 1.2 วิเคราะห์รายข้อ โดยใช้โปรแกรม CTIA แล้วคัดเลือกข้อสอบที่มีคุณภาพด้านความยากจากค่าดัชนีความยาก และคัดเลือกข้อสอบที่มีคุณภาพด้านอำนาจจำแนกจากสหสัมพันธ์ไบซีเรียล
2. วิเคราะห์ข้อมูลในขั้นตอนการวิจัย ได้ดำเนินการดังนี้
 - 2.1 ตรวจสอบให้คะแนนตามเกณฑ์ที่ตั้งไว้
 - 2.2 หาค่าความเชื่อมั่นของแบบทดสอบที่ตรวจด้วยวิธี 0-1 โดยใช้สูตร KR 20

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการหาคุณภาพของเครื่องมือที่ใช้ในการศึกษาค้นคว้า
 - 1.1 การตรวจสอบความเที่ยงตรงของแบบทดสอบวัดความถนัดทางการเรียน จากค่าดัชนีความสอดคล้อง (Index of Consistency) (บุญเชิด ภิญโญนนตพงษ์. 2545 : 95)

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

เมื่อ	IOC	แทน	ค่าดัชนีความสอดคล้อง
	$\sum R$	แทน	ผลรวมของการพิจารณาของผู้เชี่ยวชาญ
	N	แทน	จำนวนผู้เชี่ยวชาญทั้งหมด

1.2 การวิเคราะห์รายชื่อของแบบทดสอบความถนัดทางการเรียน เพื่อหาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) โดยใช้เทคนิคการแบ่งกลุ่มสูง กลุ่มต่ำ 27% แล้วใช้โปรแกรมคอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบ (CTIA) โดยมีสูตรที่เกี่ยวข้องดังนี้

1.2.1 ค่าความยากง่าย คำนวณจากดัชนีค่าความยากง่าย (Difficulty Index)

$$P = \frac{R}{N}$$

เมื่อ P แทน ดัชนีค่าความยาก

R แทน จำนวนกลุ่มตัวอย่างที่ทำข้อสอบข้อนั้นถูก

N แทน จำนวนกลุ่มตัวอย่างทั้งหมดที่ทำข้อสอบข้อนั้น

1.2.2 ค่าอำนาจจำแนก คำนวณจากสูตรสหสัมพันธ์แบบพอยท์ไบเซเรียล (Point Biserial Correlation) (บุญเชิด ภิญโญนนตพงษ์. 2521 : 257-258)

$$r_{pbis} = \frac{M_p - M_q}{S_t} \cdot \sqrt{pq}$$

เมื่อ r_{pbis} แทน ค่าอำนาจจำแนก

M_p แทน ค่าเฉลี่ยของคะแนนรวมในกลุ่มตอบถูก

M_q แทน ค่าเฉลี่ยของคะแนนรวมในกลุ่มที่ตอบผิด

S_t แทน ส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งหมด

p แทน สัดส่วนของคนตอบถูก

q แทน สัดส่วนของคนตอบผิด (1 - p)

1.3 การหาค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ก่อนและหลังการคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออก คำนวณจากสูตร KR-20 ตามวิธีของคูเดอร์-ริชาร์ดสัน (บุญเชิด ภิญโญนนตพงษ์. 2545 : 129)

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

เมื่อ r_{tt} แทน ค่าความเชื่อมั่นของแบบทดสอบ

S_t^2 แทน ความแปรปรวนของคะแนนที่สอบได้

p แทน อัตราส่วนของคนที่ทำถูกในแต่ละข้อ

q แทน 1 - p

k แทน จำนวนข้อสอบ

2. สถิติที่ใช้ในการวิจัย

2.1 สถิติที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้ วิธี

แมนเทิล-แฮนซ์เซล (Mantel-Haenszel:MH) (Camilli and Shepard.1994:116) โดยใช้สูตร ดังนี้

$$\alpha_{MH} = \frac{\sum A_j D_j / T_j}{\sum B_j C_j / T_j}$$

เมื่อ	A_j	แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบถูกในช่วงคะแนน j
	B_j	แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบผิดในช่วงคะแนน j
	C_j	แทน ผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบถูกในช่วงคะแนน j
	D_j	แทน ผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบผิดในช่วงคะแนน j
	T_j	แทน ผู้สอบทั้งสองกลุ่มที่ได้รับคะแนนรวมจากการสอบในช่วงคะแนน j
	α_{MH}	แทน สัดส่วนของการตอบข้อสอบถูกและผิดระหว่างกลุ่มในแต่ละข้อในทุกช่วงคะแนน j

2.1.1 สถิติที่ใช้ทดสอบค่าความแตกต่างจาก 1.00 อย่างมีนัยสำคัญที่ระดับชั้นแห่งความเป็นอิสระ เท่ากับ 1 ใช้สูตรดังนี้

$$\chi^2_{MH} = \frac{[\sum A_j - \sum E(A_j) - 0.5]^2}{\sum \text{Var}(A_j)}$$

เมื่อ	χ^2_{MH}	แทน ค่าการทดสอบความลำเอียงอย่างมีนัยสำคัญของข้อสอบ
	$E(A_j)$	แทน ค่าคาดหวังของ A ที่ระดับคะแนน j
	$\text{Var}(A_j)$	แทน ค่าความแปรปรวนของ A ที่ระดับคะแนน j

เมื่อ

$$E(A_j) = \frac{N_{Rj} N_{1j}}{N_j}$$

และ

$$\text{Var}(A_j) = \frac{N_{Rj} N_{Fj} N_{1j} N_{0j}}{(N_j)^2 (N_j - 1)}$$

2.2 สถิติที่ใช้ในการตรวจสอบการหาหน้าที่ต่างกันของข้อสอบโดย วิวิวิเคราะห์ความแปรปรวน (Laksana. 1979 : 12)

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}$$

เมื่อ	μ	แทน ค่าเฉลี่ยทั้งหมด
	α_i	แทน อิทธิพลจากข้อสอบข้อที่ i (i=1,2,3,...,i)
	β_j	แทน อิทธิพลจากกลุ่มผู้สอบ กลุ่มที่ j (j=1,2,3,...,j)
	$(\alpha\beta)_{ij}$	แทน ปฏิสัมพันธ์ระหว่างข้อสอบกับกลุ่มผู้สอบ
	ε_{ij}	แทน ความคลาดเคลื่อนแบบสุ่มภายในเซลล์ i,j

สถิติที่ใช้ทดสอบสมมติฐาน

2.3 สถิติทดสอบนัยสำคัญของความแตกต่างของจำนวนข้อสอบที่ทำหน้าที่ต่างกัน ใช้สูตร Q-Test (บุญเชิด ภิญโญนันตพงษ์. 2527 : 100 ; อ้างอิงจาก Hambleton, et al. 1978 : 37)

$$Q = \frac{(n-1)(n\sum p^2 - T^2)}{nT - \sum X^2}$$

เมื่อ	Q	แทน การทดสอบที่มีการแจกแจงแบบ χ^2
	n	แทน จำนวนวิวิเคราะห์ข้อสอบ
	p	แทน จำนวนข้อสอบที่ถูกคัดเลือกแต่ละวิวิเคราะห์ข้อสอบ
	X	แทน คะแนนของข้อสอบที่ถูกคัดเลือกในแต่ละข้อ
	T	แทน ผลรวมคะแนนของข้อสอบที่ถูกคัดเลือกทุกข้อ

2.4 ค่าความเชื่อมั่นของแบบทดสอบที่ปรับขยาย ด้วยสูตรของสเปียร์แมน-บราวน์ (Spearman-Brown) (บุญเชิด ภิญโญนันตพงษ์. 2545 : 123)

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

เมื่อ	r_{tt}	แทน ความเชื่อมั่นของแบบทดสอบที่ปรับขยาย
	r_{xx}	แทน ความเชื่อมั่นเดิม
	n	แทน จำนวนเท่าที่ขยายออก

2.5 สถิติทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่นหลายค่า ซึ่งประมาณจากกลุ่มตัวอย่างกลุ่มเดียว ใช้สูตร UX_1 (บุญเชิด ภิญโญอนันตพงษ์. 2537 : 57 ; อ้างอิงจาก Woodruff & Feldt. 1986 : 393-413)

$$UX_1 = \frac{\sum_{i=1}^m (u_i - \bar{u})^2}{S_u^2 - C_u}$$

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

$$u_i = \frac{1}{(1-r_i)^{1/3}}$$

$$S_u^2 = \frac{2}{9m(N_c - 1)} \sum_{i=1}^m u_i^2$$

$$C_u = \frac{4}{9m(m-1)(N_c - 1)} \sum_{i=2}^m \sum_{j=1}^{i-1} r_{ij}^2 u_i u_j$$

$$N_c = N \left(\frac{\bar{n}_h - 1}{\bar{n}_h + 1} \right)$$

$$\bar{n}_h = \frac{m}{\sum_{i=1}^m \left(\frac{1}{n_i} \right)}$$

เมื่อ UX_1 แทน สถิติทดสอบที่มีการแจกแจงแบบ χ^2 ; $df = m-1$

r แทน ค่าความเชื่อมั่นที่คำนวณได้

r_{ij} แทน ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร

m แทน จำนวนค่าสัมประสิทธิ์ความเชื่อมั่น

S_u^2 แทน ค่าความแปรปรวนของ u_i

C_u แทน ค่าความแปรปรวนร่วมของ u_i

$\bar{n}_h$ แทน ค่าเฉลี่ยฮาร์โมนิคของความยาวของแบบทดสอบ

N แทน จำนวนกลุ่มตัวอย่าง

$\bar{u}$ แทน ค่าเฉลี่ยของคะแนน

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ผู้วิจัยกำหนดสัญลักษณ์ และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

N	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
p	แทน	ค่าความยากง่ายของข้อสอบ
r	แทน	ค่าอำนาจจำแนกของข้อสอบ
K	แทน	จำนวนข้อของแบบทดสอบ
MH	แทน	การตรวจสอบการทำหน้าที่ต่างกันด้วยวิธีแมนเทิล-แฮนด์เซล
ANOVA	แทน	การตรวจสอบการทำหน้าที่ต่างกันด้วยวิธีวิเคราะห์ความแปรปรวน
N _R	แทน	จำนวนกลุ่มอ้างอิง
N _F	แทน	จำนวนกลุ่มเปรียบเทียบ
K _{diff}	แทน	จำนวนข้อสอบที่ทำหน้าที่ต่างกัน
K _{No dif}	แทน	จำนวนข้อสอบที่ทำหน้าที่ไม่ต่างกัน
Q	แทน	ค่าสถิติทดสอบที่มีการแจกแจงแบบ χ^2
r _{xx}	แทน	ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดข้อสอบที่ทำหน้าที่ต่างกันออกด้วย สูตร KR-20
r _{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบที่ปรับขยายด้วยสูตรของสเปียร์แมน-บราวน์
UX ₁	แทน	ค่าสถิติซึ่งมีการแจกแจงในรูปของ χ^2
df	แทน	ลำดับชั้นความเป็นอิสระ

การนำเสนอผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูลครั้งนี้ ผู้วิจัยได้นำเสนอตามลำดับดังนี้

1. จำนวนนักเรียนของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบ
2. การวิเคราะห์ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการต่างกัน
3. การเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน
4. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบออกแล้ว ด้วยวิธีการตรวจสอบต่างกัน

ผลการวิเคราะห์ข้อมูล

1. จำนวนนักเรียนของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบ

การวิเคราะห์ตอนนี้ผู้วิจัยได้นำคะแนนของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลขมาแจกแจงจำนวนแยกตามเกณฑ์บ่งชี้การทำหน้าที่ต่างกัน ดังนี้

กลุ่มอ้างอิง คือ เพศชาย ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ

กลุ่มเปรียบเทียบ คือ เพศหญิง ความรู้สึกไม่คุ้นเคย ความรู้สึกไม่สนใจ และความรู้สึกไม่พอใจ

ดังแสดงในตาราง 4 (และดูรายละเอียดเพิ่มเติมในตาราง 12 ภาคผนวก ค)

ตาราง 4 จำนวนนักเรียนในกลุ่มอ้างอิง กลุ่มเปรียบเทียบของแบบทดสอบวัดความถนัดทางการเรียน

แบบทดสอบวัดความถนัดทางการเรียน	จำนวนนักเรียน	
	กลุ่มอ้างอิง (คน)	กลุ่มเปรียบเทียบ (คน)
ด้านเหตุผล		
- เพศ	310	285
- ความรู้สึกคุ้นเคย	258	307
- ความรู้สึกสนใจ	469	116
- ความรู้สึกพอใจ	482	103
ด้านภาษา		
- เพศ	310	285
- ความรู้สึกคุ้นเคย	299	286
- ความรู้สึกสนใจ	491	94
- ความรู้สึกพอใจ	494	91
ด้านตัวเลข		
- เพศ	310	285
- ความรู้สึกคุ้นเคย	303	282
- ความรู้สึกสนใจ	512	73
- ความรู้สึกพอใจ	494	91

ผลการวิเคราะห์จาก ตาราง 4 พบว่า แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล มีจำนวนนักเรียนที่เป็นกลุ่มอ้างอิง ได้แก่ เพศชาย 310 คน , นักเรียนที่คุ้นเคยในข้อสอบ 258 คน , นักเรียนที่สนใจในข้อสอบ 469 คน , นักเรียนที่พอใจในข้อสอบ 482 คน และมีจำนวนนักเรียนที่เป็นกลุ่มเปรียบเทียบ ได้แก่ เพศหญิง 285 คน , นักเรียนที่ไม่คุ้นเคยในข้อสอบ 307 คน , นักเรียนที่ไม่สนใจในข้อสอบ 116 คน , นักเรียนที่ไม่พอใจในข้อสอบ 103 คน

แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา มีจำนวนนักเรียนที่เป็นกลุ่มอ้างอิง ได้แก่ เพศชาย 310 คน , นักเรียนที่คุ้นเคยในข้อสอบ 299 คน , นักเรียนที่สนใจในข้อสอบ 491 คน , นักเรียนที่พอใจในข้อสอบ 494 คน และมีจำนวนนักเรียนที่เป็นกลุ่มเปรียบเทียบ ได้แก่ เพศหญิง 285 คน , นักเรียนที่ไม่คุ้นเคยในข้อสอบ 286 คน , นักเรียนที่ไม่สนใจในข้อสอบ 94 คน , นักเรียนที่ไม่พอใจในข้อสอบ 91 คน

แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข มีจำนวนนักเรียนที่เป็นกลุ่มอ้างอิง ได้แก่ เพศชาย 310 คน , นักเรียนที่คุ้นเคยในข้อสอบ 303 คน , นักเรียนที่สนใจในข้อสอบ 512 คน , นักเรียนที่พอใจในข้อสอบ 494 คน และมีจำนวนนักเรียนที่เป็นกลุ่มเปรียบเทียบ ได้แก่ เพศหญิง 285 คน , นักเรียนที่ไม่คุ้นเคยในข้อสอบ 282 คน , นักเรียนที่ไม่สนใจในข้อสอบ 73 คน , นักเรียนที่ไม่พอใจในข้อสอบ 91 คน

2. การวิเคราะห์ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการต่างกัน

2.1 การวิเคราะห์ตอนนี้ผู้วิจัยได้นำคะแนนของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบมาวิเคราะห์การทำหน้าที่ต่างกันของแบบทดสอบ ดังแสดงในตาราง 5

ตาราง 5 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้วยวิธีการตรวจสอบต่างกัน

แบบทดสอบวัด		ข้อที่การทำหน้าที่ต่างกัน	
ความถนัดทางการ	$N_R : N_F$	วิธี MH	วิธี ANOVA
เรียน / กลุ่มอ้างอิง			
ด้านเหตุผล			
- เพศ	310 : 285	7 ข้อ (ร้อยละ 23.33) 7, 10, 18, 19, 20, 27, 29	17 ข้อ (ร้อยละ 56.67) 2, 3, 4, 7, 10, 11, 12, 13, 17, 20, 21, 22 , 23, 24, 26, 27, 29
- ความรู้สึกคุ้นเคย	258 : 307	1 ข้อ (ร้อยละ 3.33) 3	5 ข้อ (ร้อยละ 16.67) 3, 15, 19, 25, 29
- ความรู้สึกสนใจ	469 : 116	0 ข้อ	8 ข้อ (ร้อยละ 26.67) 3, 10, 11, 13, 17, 20, 22, 30
- ความรู้สึกพอใจ	482 : 103	2 ข้อ (ร้อยละ 6.67) 17, 26	3 ข้อ (ร้อยละ 10.00) 17, 26, 30

ผลการวิเคราะห์จาก ตาราง 5 พบว่า แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นเพศ ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 7 ข้อ ได้แก่ ข้อที่ 7, 10, 18, 19, 20, 27, 29 คิดเป็นร้อยละ 23.33 ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 17 ข้อ ได้แก่ ข้อที่ 2, 3, 4, 7, 10, 11, 12, 13, 17, 20, 21, 22, 23, 24, 26, 27, 29 คิดเป็นร้อยละ 56.67 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 5 ข้อ ได้แก่ ข้อที่ 7, 10, 20, 27, 29

จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย และตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 1 ข้อ ได้แก่ ข้อที่ 3 คิดเป็นร้อยละ 3.33 ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 5 ข้อ ได้แก่ ข้อที่ 3, 15, 19, 25, 29 คิดเป็นร้อยละ 16.67 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 1 ข้อ ได้แก่ ข้อที่ 3

จากกลุ่มอ้างอิงที่เป็นความรู้สึกสนใจเป็นกลุ่มอ้างอิง ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล ไม่พบข้อสอบที่ทำหน้าที่ต่างกัน ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 8 ข้อ ได้แก่ ข้อที่ 3, 10, 11, 13, 17, 20, 22, 30 คิดเป็นร้อยละ 26.67

จากกลุ่มอ้างอิงที่เป็นความรู้สึกพอใจ ตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เซล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 17, 26 คิดเป็นร้อยละ 6.67 ตรวจสอบด้วย วิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 3 ข้อ ได้แก่ ข้อที่ 17, 26 , 30 คิดเป็นร้อยละ 10.00 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 2 ข้อ ได้แก่ ข้อที่ 17, 26

2.2 การวิเคราะห์ตอนนี้อยู่ภายใต้คำแนะนำของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบมาวิเคราะห์การทำหน้าที่ต่างกันของแบบทดสอบ ดังแสดงในตาราง 6

ตาราง 6 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ด้วยวิธีการตรวจสอบต่างกัน

แบบทดสอบวัด		ข้อที่การทำหน้าที่ต่างกัน	
ความถนัดทางการ	$N_R : N_F$	วิธี MH	วิธี ANOVA
เรียน / กลุ่มอ้างอิง			
ด้านภาษา			
- เพศ	310 : 285	9 ข้อ (ร้อยละ 30.00) 34, 35, 40, 42, 44, 45, 52, 54, 55	10 ข้อ (ร้อยละ 33.39) 39, 44, 45 , 47, 50, 51, 52, 54 , 55 , 59
- ความรู้สึกคุ้นเคย	299 : 286	0 ข้อ	0 ข้อ
- ความรู้สึกสนใจ	491 : 94	3 ข้อ (ร้อยละ 10.00) 33, 43, 55	2 ข้อ (ร้อยละ 6.67) 43,55
- ความรู้สึกพอใจ	494 : 91	2 ข้อ (ร้อยละ 6.67) 42, 59	2 ข้อ (ร้อยละ 6.67) 54, 59

ผลการวิเคราะห์จากตาราง 6 พบว่า แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากกลุ่มอ้างอิงที่เป็นเพศ ตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เซล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 9 ข้อ ได้แก่ ข้อที่ 34, 35, 40, 42, 44, 45, 52, 54, 55 คิดเป็นร้อยละ 30.00 ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 10 ข้อ ได้แก่ ข้อที่ 39, 44, 45, 47, 50, 51, 52, 54, 55, 59 คิดเป็นร้อยละ 33.39 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธีการตรวจสอบ จำนวน 5 ข้อ ได้แก่ ข้อที่ 44, 45, 52, 54, 55

จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ทั้ง 2 วิธีการตรวจสอบไม่พบข้อสอบที่ทำหน้าที่ต่างกัน

จากกลุ่มอ้างอิงที่เป็นความรู้สึกสนใจ ตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เซล พบข้อสอบที่ทำหน้าที่ต่างกันรวม 3 ข้อ ได้แก่ ข้อที่ 33, 43, 55 คิดเป็นร้อยละ 10.00 ตรวจสอบด้วย วิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 43, 55คิดเป็นร้อยละ 6.67 และมีจำนวนข้อสอบ ที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 2 ข้อ ได้แก่ ข้อที่ 43, 55

จากกลุ่มอ้างอิงที่เป็นความรู้สึกพอใจ ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 42, 59 คิดเป็นร้อยละ 6.67 ตรวจสอบด้วย วิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 54, 59 คิดเป็นร้อยละ 6.67 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 1 ข้อ ได้แก่ ข้อที่ 59

2.3 การวิเคราะห์ตอนนี้ผู้วิจัยได้นำคะแนนของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบมาวิเคราะห์การทำหน้าที่ต่างกันของแบบทดสอบ ดังแสดงในตาราง 7

ตาราง 7 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน

แบบทดสอบวัด		ข้อที่การทำหน้าที่ต่างกัน	
ความถนัดทางการ	$N_R : N_F$	วิธี MH	วิธี ANOVA
เรียน / กลุ่มอ้างอิง			
<u>ด้านตัวเลข</u>			
- เพศ	310 : 285	4 ข้อ (ร้อยละ 13.33) 71, 79, 82, 86	4 ข้อ (ร้อยละ 13.33) 63, 79, 82, 86
- ความรู้สึกคุ้นเคย	299 : 286	3 ข้อ (ร้อยละ 10.00) 65, 70, 80	2 ข้อ (ร้อยละ 6.67) 70, 80
- ความรู้สึกสนใจ	491 : 94	3 ข้อ (ร้อยละ 10.00) 69, 77, 78	3 ข้อ (ร้อยละ 10.00) 69, 77, 78
- ความรู้สึกพอใจ	494 : 91	0 ข้อ	0 ข้อ

ผลการวิเคราะห์จากตาราง 7 พบว่า แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากกลุ่มอ้างอิงที่เป็นเพศ ตรวจสอบด้วย วิธีแมนเทิล-แฮนดเชล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 4 ข้อ ได้แก่ ข้อที่ 71, 79, 82, 86 คิดเป็นร้อยละ 13.33 ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 4 ข้อ ได้แก่ ข้อที่ 63, 79, 82, 86 คิดเป็นร้อยละ 13.33 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 3 ข้อ ได้แก่ ข้อที่ 79, 82, 86

จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 42, 59 คิดเป็นร้อยละ 6.67 ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน พบจำนวนข้อสอบที่ทำหน้าที่ต่างกันรวม 2 ข้อ ได้แก่ ข้อที่ 54, 59 คิดเป็นร้อยละ 6.67 และมีจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี จำนวน 1 ข้อ ได้แก่ ข้อที่ 59

จากกลุ่มอ้างอิงที่เป็นความรู้สึกสนใจ ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชลและวิธีวิเคราะห์ความแปรปรวนพบจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้ง 2 วิธี รวม 3 ข้อ ได้แก่ ข้อที่ 69, 77, 78 คิดเป็นร้อยละ 10.00

จากกลุ่มอ้างอิงที่เป็นความรู้สึกพอใจ ทั้ง 2 วิธีการตรวจสอบไม่พบข้อสอบที่ทำหน้าที่ต่างกัน

3. การเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน

การวิเคราะห์ตอนนี้ผู้วิจัยนำจำนวนข้อที่ทำหน้าที่ต่างกันของข้อสอบระหว่าง เพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ ด้วยวิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน มาเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกัน โดยใช้สถิติ Q- Test ปรากฏผลทดสอบความแตกต่างอย่างมีนัยสำคัญทางสถิติ ดังแสดงในตาราง 8

ตาราง 8 การเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน

แบบทดสอบวัดความถนัดทาง การเรียน / กลุ่มอ้างอิง	วิธี MH		วิธี ANOVA		Q
	K _{dif}	K _{No dif}	K _{dif}	K _{No dif}	
<u>ด้านเหตุผล</u>					
- เพศ	7	23	17	13	7.1429*
- ความรู้สึกคุ้นเคย	1	29	5	25	3.0000
- ความรู้สึกสนใจ	0	30	8	22	8.0000*
- ความรู้สึกพอใจ	2	28	3	27	1.0000
<u>ด้านภาษา</u>					
- เพศ	9	21	10	20	0.1111
- ความรู้สึกคุ้นเคย	0	30	0	30	0.0000
- ความรู้สึกสนใจ	3	27	2	28	1.0000
- ความรู้สึกพอใจ	2	28	2	28	0.0000
<u>ด้านตัวเลข</u>					
- เพศ	4	26	4	26	0.0000
- ความรู้สึกคุ้นเคย	3	27	2	28	1.0000
- ความรู้สึกสนใจ	3	27	3	27	0.0000
- ความรู้สึกพอใจ	0	30	0	30	0.0000

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์จากตาราง 6 พบว่า จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้วยวิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกันจากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ ได้จำนวนข้อสอบแตกต่างกันอย่างมีนัยสำคัญ

ทางสถิติที่ระดับ .05 จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา และด้านตัวเลข ด้วยวิธีแมนเทิล-แฮนด์เซล และวิธีวิเคราะห์ความแปรปรวน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกัน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบออกแล้ว ด้วยวิธีการตรวจสอบต่างกัน

4.1 การวิเคราะห์ตอนนี้ผู้วิจัยได้นำแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข ที่คัดข้อสอบที่ทำหน้าที่ต่างกัน ด้วยวิธีแมนเทิล-แฮนด์เซล และวิธีวิเคราะห์ความแปรปรวน ออกไป แล้วนำจำนวนข้อสอบที่ใช้ได้มาคำนวณค่าความเชื่อมั่นของแบบทดสอบ (r_{xx}) ด้วยสูตร KR-20 แล้วจึงปรับขยายค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านให้มีจำนวนข้อเท่ากับ 30 ข้อ (r_{tt}) ด้วยสูตรสเปียร์แมน-บราวน์ (Spearman-Brown) ดังแสดงในตาราง 9

ตาราง 9 ค่าความเชื่อมั่นของแบบทดสอบหลังการปรับขยายให้แบบทดสอบมีจำนวนข้อเท่ากับ 30 ข้อ

แบบทดสอบวัดความ ถนัดทางการเรียน / กลุ่มอ้างอิง	วิธี MH			วิธี ANOVA		
	จำนวนข้อ สอบที่ใช้ได้	r_{xx}	r_{tt}	จำนวนข้อ สอบที่ใช้ได้	r_{xx}	r_{tt}
ด้านเหตุผล						
- เพศ	23	0.7079	0.7597	13	0.4583	0.9269
- ความรู้สึกคุ้นเคย	29	0.7442	0.7699	25	0.7274	0.7620
- ความรู้สึกสนใจ	30	0.7497	0.7497	22	0.6575	0.8966
- ความรู้สึกพอใจ	28	0.7288	0.7422	27	0.7289	0.7492
ด้านภาษา						
- เพศ	21	0.6369	0.7148	20	0.6459	0.7324
- ความรู้สึกคุ้นเคย	30	0.7660	0.7660	30	0.7660	0.7660
- ความรู้สึกสนใจ	27	0.7436	0.8262	28	0.7591	0.7715
- ความรู้สึกพอใจ	28	0.7443	0.7572	28	0.7457	0.7586
ด้านตัวเลข						
- เพศ	26	0.8327	0.8517	26	0.8321	0.8512
- ความรู้สึกคุ้นเคย	27	0.8363	0.8502	28	0.8443	0.8532
- ความรู้สึกสนใจ	27	0.8343	0.8484	27	0.8343	0.8484
- ความรู้สึกพอใจ	30	0.8531	0.8531	30	0.8531	0.8531

ผลการวิเคราะห์จากตาราง 10 พบว่า ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เชล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 23 , 29 , 30 และ 28 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.7079 , 0.7442 , 0.7497 และ 0.7288 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{tt}) เท่ากับ 0.7597 , 0.7699 , 0.7497 และ 0.7422 ตามลำดับ ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 13 , 25 , 22 และ 27 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.4583 , 0.7274 , 0.6575 และ 0.7289 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{tt}) เท่ากับ 0.9269 , 0.7620 , 0.8966 และ 0.7492 ตามลำดับ

ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ตรวจสอบด้วยวิธีแมนเทิล-แฮนด์เชล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 21 , 30 , 27 และ 28 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.6369 , 0.7660 , 0.7436 และ 0.7443 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{tt}) เท่ากับ 0.7148 ,

0.7660 , 0.8262 และ 0.7572 ตามลำดับ ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 20 , 30 , 28 และ 28 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.6459 , 0.7660 , 0.7591 และ 0.7457 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{11}) เท่ากับ 0.7324 , 0.7660 , 0.7715 และ 0.7586 ตามลำดับ

ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ตรวจสอบด้วยวิธีแมนเทิล-แฮนดเชล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 26 , 27 , 27 และ 30 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.8327 , 0.83632 , 0.8343 และ 0.8531 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{11}) เท่ากับ 0.8517 , 0.8502 , 0.8484 และ 0.8531 ตามลำดับ ตรวจสอบด้วยวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ มีจำนวนข้อสอบที่ใช้ได้เท่ากับ 26 , 28 , 27 และ 30 ตามลำดับ มีค่าความเชื่อมั่น (r_{xx}) เท่ากับ 0.8321 , 0.8443 , 0.8343 และ 0.8531 ตามลำดับ เมื่อนำมาปรับขยายให้มีจำนวนเท่ากับ 30 ข้อ มีค่าความเชื่อมั่น (r_{11}) เท่ากับ 0.8512 , 0.8532 , 0.8484 และ 0.8531 ตามลำดับ

4.2 การวิเคราะห์ตอนนี้ผู้วิจัยนำค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา ด้านตัวเลข ที่ปรับขยายเป็น 30 ข้อ มาเปรียบเทียบกับค่าความเชื่อมั่นของแบบทดสอบด้วยสูตร UX, ดังแสดงในตาราง 10

ตาราง 10 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้วย UX₁

แบบทดสอบวัดความถนัด ทางการเรียน / กลุ่มอ้างอิง	จำนวนข้อ (K)	ค่าความเชื่อมั่น r_{tt}		UX ₁
		วิธี MH	วิธี ANOVA	
<u>ด้านเหตุผล</u>				
- เพศ	30	0.7597	0.9269	181.4901*
- ความรู้สึกคุ้นเคย	30	0.7699	0.7620	0.1556
- ความรู้สึกสนใจ	30	0.7497	0.8966	103.0023*
- ความรู้สึกพอใจ	30	0.7422	0.7492	0.1035
<u>ด้านภาษา</u>				
- เพศ	30	0.7148	0.7324	0.5539
- ความรู้สึกคุ้นเคย	30	0.7660	0.7660	0.0000
- ความรู้สึกสนใจ	30	0.8262	0.7715	10.1898*
- ความรู้สึกพอใจ	30	0.7572	0.7586	0.0046
<u>ด้านตัวเลข</u>				
- เพศ	30	0.8517	0.8512	0.0016
- ความรู้สึกคุ้นเคย	30	0.8502	0.8532	0.5590
- ความรู้สึกสนใจ	30	0.8484	0.8484	0.0000
- ความรู้สึกพอใจ	30	0.8531	0.8531	0.0000

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์จากตาราง 10 พบว่า ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้วยวิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ด้วยวิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จากกลุ่มอ้างอิงที่เป็น ความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้วยวิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน จากกลุ่มอ้างอิงที่เป็น เพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

จุดมุ่งหมายของการศึกษาค้นคว้า

การศึกษาในครั้งนี้มีจุดมุ่งหมายสำคัญ เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา ด้านตัวเลข ตามความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ ด้วยวิธีการตรวจสอบต่างกัน 2 วิธี คือ วิธี แมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน เป็นจุดมุ่งหมายเฉพาะของการศึกษาค้นคว้า ต่อไปนี้

1. เพื่อศึกษาจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน
2. เปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา ด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน
3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา ด้านตัวเลข หลังคัดเลือกข้อที่ทำหน้าที่ต่างกันออก ด้วยวิธีการตรวจสอบต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนสังกัดกรุงเทพมหานคร กลุ่มรัตนโกสินทร์ ซึ่งได้มาโดยการสุ่มแบบ 2 ขั้นตอน (Two-Stage Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) โรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) จำนวนโรงเรียน 9 โรงเรียน จำนวน 22 ห้องเรียน รวมนักเรียนทั้งสิ้น 585 คน จำแนกเป็นนักเรียนชาย 310 คน นักเรียนหญิง 275 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดความถนัดทางการเรียนที่สร้างโดยยึดหลักทฤษฎีหลายองค์ประกอบ ตามแนวของเทอร์สโตน (Thurstone) ดังนี้

1. แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล (R) (ข้อที่ 1 ถึง 30)
2. แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา (V) (ข้อที่ 31 ถึง 60)
3. แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข (N) (ข้อที่ 61 ถึง 90)

เป็นข้อสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก จำนวน 90 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง

วิธีเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อนำมาวิเคราะห์นั้น ผู้วิจัยดำเนินการตามขั้นตอน ดังนี้

1. กำหนดเวลาที่จะทำการทดสอบ โดยติดต่อขออนุญาตผู้บริหารโรงเรียนที่เลือกเป็นกลุ่มตัวอย่าง
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่จะทำการสอบในแต่ละครั้ง วางแผนดำเนินการสอบ
3. ชี้แจงทำความเข้าใจกับผู้ประสานในแต่ละโรงเรียน ในเรื่องลักษณะของแบบทดสอบ และวิธีดำเนินการสอบ

4. นำแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา และด้านตัวเลข ไปทดสอบกับกลุ่มตัวอย่างจำนวน 585 คน เพื่อศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน และทำการทดสอบสมมติฐานการศึกษาค้นคว้า

การนำเสนอผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูลครั้งนี้ ผู้วิจัยได้นำเสนอตามลำดับดังนี้

1. จำนวนนักเรียนของกลุ่มอ้างอิง และกลุ่มเปรียบเทียบ
2. การวิเคราะห์ผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการต่างกัน
3. การเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้านภาษา และด้านตัวเลข ด้วยวิธีการตรวจสอบต่างกัน
4. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบออกแล้ว ด้วยวิธีการตรวจสอบต่างกัน

สรุปผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูล สรุปได้ดังนี้

1. จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้วยวิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกันจากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ ได้จำนวนข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา และด้านตัวเลข ด้วยวิธีแมนเทิล-แฮนดเชล และวิธีวิเคราะห์ความแปรปรวน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกัน จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจ ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกโดยใช้วิธีการตรวจสอบต่างกัน พบว่า

2.1 แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2.2 แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จากกลุ่มอ้างอิงที่เป็น ความรู้สึกสนใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2.3 แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากกลุ่มอ้างอิงที่เป็น เพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ ความรู้สึกพอใจ มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

จากการศึกษาครั้งนี้ สามารถอภิปรายผลได้ดังนี้

1. จำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ด้วยวิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน สามารถบ่งชี้ข้อที่ทำหน้าที่ต่างกันจากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ ได้จำนวนข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งผลที่ได้สอดคล้องกับสมมติฐานการวิจัย และยังสอดคล้องกับงานวิจัยของ เจน จันทรแสง (2537:87) ที่ศึกษาการทำหน้าที่ต่างกันของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 พบว่า กลุ่มอ้างอิงที่เป็นเพศ มีผลต่อการทำหน้าที่ต่างกันของข้อสอบ ส่วนจำนวนข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ และแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา และด้านตัวเลข ได้จำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐาน ทั้งนี้อาจเป็นเพราะ จำนวนข้อสอบที่ทำหน้าที่ต่างกันมีจำนวนข้อใกล้เคียงกัน และคะแนนผลการสอบมีความแตกต่างในสัดส่วนของการตอบข้อสอบระหว่างผู้สอบที่มีความสามารถระดับเดียวกัน และต่างระดับกัน ทำให้วิธีวิเคราะห์ความแปรปรวน มีความไวในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อมีความแตกต่างของค่าความยากของกลุ่มผู้สอบสองกลุ่ม (Angoff อ้างใน Berk 1982 : 107) และยังสอดคล้องกับงานวิจัยของ วิภา จำมัน (2544:52) ที่ศึกษาเปรียบเทียบผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันของแบบทดสอบวัดความสามารถด้านภาษา ตามระดับความง่ายของข้อสอบ ด้วยวิธีแมนเทิล-แฮนด์เชล และวิธีชิปเทสต์ พบว่าจำนวนข้อสอบที่ทำหน้าที่ต่างกันที่พบระหว่างกลุ่มที่มีระดับความง่ายต่ำ ระดับความง่ายปานกลาง ระดับความง่ายสูง มีจำนวนข้อไม่แตกต่างกันมาก เมื่อนำมาทดสอบนัยสำคัญทางสถิติ จึงมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล ด้านภาษา และด้านตัวเลข หลังจากคัดเลือกข้อสอบที่ทำหน้าที่ต่างกันออกโดยใช้วิธีการตรวจสอบต่างกัน ทำการปรับขยายค่าความเชื่อมั่นของแบบทดสอบให้มีจำนวนข้อเท่ากันด้วยสูตรสเปียร์แมน-บราวน์ พบว่า จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกสนใจ และแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากกลุ่มอ้างอิงที่เป็นความรู้สึกสนใจ พบว่าค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งผลที่ได้สอดคล้องกับสมมติฐานการวิจัย ทั้งนี้เนื่องจากการคัดข้อสอบที่ทำหน้าที่ต่างกันออกเป็นข้อที่ไม่ซ้ำกันทุกข้อ และมีจำนวนข้อไม่เท่ากัน ทำให้ค่าความเชื่อมั่นของแบบต่างกัน ดังที่บุญเชิด ภิญโญนนตพงษ์ (2521 : 313) กล่าวว่า จำนวนข้อของแบบทดสอบมีอิทธิพลต่อค่าความเชื่อมั่น แบบทดสอบใดมีจำนวนข้อน้อยจะมีค่าความเชื่อมั่นต่ำ ถ้ามีจำนวนข้อมากจะมีค่าความเชื่อมั่นสูง และยังสอดคล้องกับงานวิจัยของ พรรณี จินตมาศ (2540:86) ที่ศึกษาเปรียบเทียบผลการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแปลงค่าความยากและวิธีชิปเทสต์ในกลุ่มผู้สอบ 500 คน พบว่าค่าความเชื่อมั่นของแบบทดสอบหลังจากที่คัดข้อสอบที่ทำหน้าที่ต่างกันออกแล้วมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มอ้างอิงที่เป็นความรู้สึกคุ้นเคย ความรู้สึกพอใจ และแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา และด้านตัวเลข จากกลุ่มอ้างอิงที่เป็นเพศ ความรู้สึกคุ้นเคย ความรู้สึกสนใจ ความรู้สึกพอใจ พบว่าค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานการวิจัย ทั้งนี้เนื่องจากการคัดข้อสอบที่ทำหน้าที่ต่างกันออกนั้น พบว่าเป็นข้อสอบที่ซ้ำกัน มากกว่าแต่สอดคล้องกับงานวิจัยของ นวลอนงค์ นันทวงศ์ (2543:69) ที่ศึกษาเปรียบเทียบค่าความเชื่อมั่นหลังจากคัดข้อที่ทำหน้าที่ต่างกันออกแล้ว ระหว่างวิธีโคสแควร์ วิธีแมนเทิล-แฮนด์เชล และวิธีวิเคราะห์ความแปรปรวน พบว่ามีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ข้อเสนอแนะ

ผลของการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้สรุปข้อเสนอแนะไว้ ดังนี้

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 ควรกำหนดให้การศึกษาการทำหน้าที่ต่างกันของข้อสอบเป็นขั้นตอนหนึ่งในการสร้างแบบทดสอบ เพื่อผลการทดสอบมีความยุติธรรมมากขึ้น และคุณภาพของแบบทดสอบดีขึ้น

1.2 ข้อสอบที่ทำหน้าที่ต่างกันที่ได้จากการตรวจสอบการทำหน้าที่ต่างกันโดยมี ความรู้สึกคุ้นเคย ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ เป็นกลุ่มอ้างอิงนั้น ไม่ได้หมายความว่าข้อสอบนั้นขาดคุณภาพทางการวัด แต่เป็นการเสนอเพื่อปรับปรุง แก้ไขข้อสอบให้ดีขึ้น ทั้งนี้ในการทำข้อสอบไปใช้ขึ้นอยู่กับสถานะของผู้สอบในขณะนั้นด้วย

2. ข้อเสนอแนะในการทำวิจัย

2.1 ควรมีการศึกษาเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบอื่นๆ เช่น แบบทดสอบวัดความถนัดด้านอื่น แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน แบบทดสอบวัดเจตคติ และเปรียบเทียบกับตัวแปรด้านอื่นๆ เช่น ระดับผลการเรียน พื้นฐานทางครอบครัว ระดับชั้นการศึกษา สภาพทางภูมิศาสตร์ต่างๆ

2.2 ควรมีการศึกษาในทำนองเดียวกันนี้โดยใช้วิธีการตรวจสอบวิธีอื่นๆ

บรรณานุกรม

บรรณานุกรม

- คมศักดิ์ ชื่นชม. (2539). การศึกษาความลำเอียงที่ใช้วิธีต่างกันของแบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- จิตสุตา ธารพร. (2539). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์และแบบทดสอบความเข้าใจทางภาษา. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- เจน จันทรแสง. (2537). การวิเคราะห์ความลำเอียงของข้อสอบของแบบสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา) พิษณุโลก: บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร. ถ่ายเอกสาร.
- ชวาล แพรัตกุล. (2517). การทดสอบเพื่อค้นและพัฒนา. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- ชาญวิทย์ เทียมบุญประเสริฐ. (2528). การวัดความถนัด. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- ญาณภัทร สีหะมงคล. (2540). การเปรียบเทียบความสอดคล้องของผลการตรวจข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Raju's Area Measures และวิธี Close Interval Area. ปรินญาณพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- นิรมล ชัยขวลิต. (2537). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของแบบทดสอบความเข้าใจในการอ่านภาษาไทยตามทฤษฎีคลาสสิกคอลที่ใช้วิธีต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- บุญเชิด ภิญโญอนันตพงษ์. (2521). การวัดและประเมินผลการศึกษา. กรุงเทพฯ : ภาควิชาพื้นฐานทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- (2527). การทดสอบแบบอิงเกณฑ์ : แนวคิดและวิธีการ. กรุงเทพฯ : ภาควิชาพื้นฐานทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์. ปรินญาณพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- (2545). “ประมวลสาระชุดวิชา การพัฒนาเครื่องมือสำหรับการประเมินการศึกษา (หน่วยที่ 3). กรุงเทพฯ : มหาวิทยาลัยสุโขทัยธรรมาธิราช สาขาวิชาศึกษาศาสตร์.
- บุญส่ง นิลแก้ว. (2519). การวัดผลทางจิตวิทยา. กรุงเทพฯ : แพร์พิตยา.
- พรรณี จินตมาศ. (2540). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้ขนาดกลุ่มผู้สอบและวิธีวิเคราะห์ต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- พัชรีย์ ปิยะภักดิ์. (2531). การวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
ชั้นประถมศึกษาปีที่ 6. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิต
วิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พิศมัย สาระกุล. (2542). การสร้างแบบทดสอบวัดความสามารถทางสมองตามทฤษฎีเชาวันปัญญาของ
สเติร์นเบอร์ก (Triarchic Theory). ถ่ายเอกสาร. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา).
กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- เรวดี อินทสระ. (2539). ผลการตรวจสอบความลำเอียงของข้อสอบต่อการศึกษาความเที่ยงตรงเชิง
พยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน. ปรินญานิพนธ์ กศ.ด. (การทดสอบ
และวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่าย
เอกสาร.
- ล้วน สายยศ และ อังคนา สายยศ. (2538). เทคนิคการวิจัยทางการศึกษา. พิมพ์ครั้งที่ 4. กรุงเทพฯ: สุวีริยา
สาสน์.
- (2541). เทคนิคการสร้างและสอบข้อสอบความถนัดทางการเรียน. พิมพ์
ครั้งที่ 3. กรุงเทพฯ: ชมรมเด็ก.
- วิภา จำนั. (2544). การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถ
ด้านภาษา ตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน. ถ่ายเอกสาร. ปรินญานิพนธ์
กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
ถ่ายเอกสาร
- สมศักดิ์ จันทอง. (2541, พฤษภาคม-สิงหาคม). "จากความลำเอียงของข้อสอบ(Item Bias) มาเป็นการทำ
หน้าที่ต่างกันของข้อสอบ(DIF)," วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร. 3(1) : 81-93.
- อังคนา สายยศ. (2539). การศึกษาผลการวิเคราะห์ความลำเอียงของแบบทดสอบความเข้าใจในการอ่าน
ตามลักษณะการใช้ภาษาพูดในครอบครัว. กรุงเทพฯ: ภาควิชาการวัดผลและวิจัยทางการ
ศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- อารี สันหนวี. (2535). พหุปัญญาและการเรียนแบบร่วมมือ. กรุงเทพฯ : สมาคมเพื่อการศึกษาเด็ก.
- อารี วัชรโสติกุล. (2543). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้รูปแบบ
และวิธีการต่างกัน. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- Allen, M.J. & Yen, W.M. (1979). *Introduction to Measurement Theory*. Monterey, CA : Brooks/Cole.
- Allaof, A.V., Hambleton, R.K. & Sireci, S.G. (1999, Fall). "Identifying the Causes of DIF in Translated
Verbal Items," *Journal of Educational Measurement*. 36(3) : 185-199.
- Anatasi, Anne. (1988). *Psychological Testing*. 6th ed. New York : Macmillan.
- Berk, Ronald A. (1982). *Handbook of Method for Detecting Test Bias*. John Hopkins University Press.
- Bingham, W.V. (1937). *Aptitudes and Aptitude Testing*. New York : Harper and Brothers.
- Camilli, G. & Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. Thousand Oaks, Ca :
SAGE.

- Cang, Hua-Hua, Mazzeo, J. & Roussos, L. (1996, Fall). "Detecting DIF for Polytomously Scored Items : An Adaptation of the SIBTES Procedure," *Journal of Education Measurement*. 33(3) : 333-335.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ : Erlbaum.
- Cronbach. (1970). *Essentials of Psychological Testing*. 3rd ed. New York : Harper & Row.
- Doolittle, Allen E. & Cleary, Anne. T. (1987, Summer). "Gender-Based Differential Item Performance in Mathematics Achievement Items," *Journal of Educational Measurement*. 24(2) : 157-166.
- Hambleton, Ronald K. & Roger H.J. (1989, December). "Detecting Potentially Biased Test Item: Comparison of IRT Area and Mantel-Haenszel Methods," *Applied Psychological Measurement*. 2(4) : 313-334.
- Holland, P.W. & Wainer, H. (1993). *Differential Item Functioning*. New Jersey : Lawrence Erlbaum
- Holland, P.W. & Thayer, D.T. (1988). *Differential item Performance and the Mantel-Haenszel Procedure*. In Wainer, Howard and Brgun, Henty / *Test Validity*. Hillsdale, NJ : Erlbaum.
- Intasuwan, P. (1979). *A Comparison of Three Approaches for Determining Item Bias in Cross-National Testing*. Dissertation Ph.D.(Educational Research). Pittsburgh : Graduate School University of Pittsburgh. Photocopied.
- Kelderman, H. (1989). "Item Bias Detection Using Loglinear IRT," *Psychometrika*. 54(5) : 681-697.
- Kelderman, H. & Georerdy, B.M. (1990, Winter). "The Use of loglinear Models for Assessing Differential Item Functioning Across Manifest and Latent Examinee Groups," *Journal of Educational Measurement*. 27(4) : 370-427.
- Kim, Seock-Ho. & Cohen, Allan S. (1991, September). "A Comparison of Two Area Measures for Detecting Differential Item Functioning," *Applied Psychological Measurement*. 15(3) : 269-278.
- Kirk, R.E. (1995). *Experimental Design : Procedures for the behavioral sciences*. 3rd ed. Pacific Grove, CA : Brooks/Cole.
- Laksana, S. and W.E., Coffman. (1980). *A Comparison of ANOVA Approach and ICC Approach for Assessing Item Bias in an Achievement Test*. Iowa Testing Programs Occasional paper. The University of Iowa.
- Mazor, K.M., Clauser, B.E. & Hambleton, R.K. (1992, Summer). "The effect of sample size on the functioning of the Mantel-Haenszel Statistic," *Educational and Psychological Measurement*. 52(2) : 443-451.
- Mellenbergh, G. (1982, Summer). "Contingency Table Models for Assessing Item Bias," *Journal of Educational Statistics*. 7(2) : 105-118.
- Millsap, R.E. & Everson, H.T. (1993, December). "Methodology Review : Statistical Approaches for Assessing Measurement Bias," *Applied Psychological Measurement*. 17(4) : 297-334.

- Narayanan, P. & Swaminathan, H. (1994,December). "Performance of the Mantel-Haenszel and Simultaneous item Bias Procedures for Detecting Differential Item Functioning," *Applied Psychological Measurement*. 18(4) : 315-328.
- Oshima,T.C. & Miller, M.D. (1992,September). "Multidimensionality and item Bias in Item Response Theory," *Applied Psychological Measurement*. 16(3) : 237-248.
- Potenza, Maria T. & Dorans, Neil J. (1995,March). "DIF Assessment for Polytomously Scored Items : A Framework for Classification and Evaluation," *Applied Psychological Measurement*. 19(1) : 23-27.
- Raju, Drasgow,F. & Slinde. (1993,Summer)."An Empirical Comparison of the Area Methods, Lord's Chi-Square Test, and the Mantel-Haenszel Technique for Assessing Differential Item Functioning," *Educational and Psychological Measurement*. 53(2) : 301-314.
- Rogers, H.J. & Swaminathan, H. (1993,June). "A Comparision of Logistic regression and Mantel-Haenszel Procedures for Detecting Differential Item Functioning," *Applied Psychological Measurement*. 17(2) : 105-116.
- Roussos, Louis A. & Stout, William F. (1996,Summer). "Simulation Studied of the Effect of Small Sample Size and Studied Item Parameter on SIBTEST and Mantel-Haenszel Type I Error Performance," *Journal of Educational Measurement*. 33(2) : 215-230.
- Sax,Guilbert. (1989). *Principle of Educational and Psychological Measurement and Evaluation*. 3rd ed. Belmont,CA : Wadsworth.
- Scheuneman, J.D. (1979,Summer). "A Method for Assessing Bias in Test Items," *Journal of Educational Measurement*. 16(2) : 143-152.
- Shealy, Robin. & Stout, William. (1993,June). "A Model-Based Standardization Approach that Separates True Bias/DIF Form Group Ability Differences and Detects Test Bias/DTF as well as Item Bias/DIF," *Psychometrika*. 58(2) : 159-194.
- Shepard, L.A. Camilli, G. & Averill, M. (1981,Winter) . "test Item Bias," *Journal of Educational Statistics*. 6(4) : 371-375.
- Stricker, L. & Emmerich, W. (1999,Winter). "Possible Determinants of Differential Item Functioning : Familiarity Interest and Emotional Reaction," *Journal of Educational Measurement*. 36(4) : 347-366.
- Swaminathan, H. & Rogers, H.J. (1990,Winter). "Detecting Differential Item Functioning Using Logistic Regression Procedures," *Journal of Educational Measurement*. 27(4) : 361-370.
- Thissen,D., Steinberg, L. & Wainer, H. (1988). *Use of Item Response Theory in the Study of Group Differences in Trace Lines* . In H. Wainer & H. Braum (Eds.), *Test Validity*. Hillsdale NJ : Erlbaum.
- Warren,H.C. (1934). *Dictionary of Psychology*. New York : Houghton Mifflin.

ภาคผนวก

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญตรวจสอบแนะนำแก้ไขเครื่องมือ

รายชื่อผู้เชี่ยวชาญ

รศ. นิภา ศรีไพโรจน์	ภาควิชาวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
รศ. ชูศรี วงศ์รัตนะ	ภาควิชาวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
รศ. วัณญา วิศาลาภรณ์	ภาควิชาวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ผศ. ระวีวรรณ พันธุ์พานิช	ภาควิชาวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อ.ดร. สุพพร เข้มแข็ง	ภาควิชาวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

ภาคผนวก ข

ตารางแสดงค่าความยากง่าย (p) และค่าอำนาจจำแนก (r)
ของแบบทดสอบวัดความถนัดทางการเรียน

ตาราง 11 ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดความถนัดทางการเรียน

ด้านเหตุผล			ด้านภาษา			ด้านตัวเลข		
ข้อที่	ค่า p	ค่า r	ข้อที่	ค่า p	ค่า r	ข้อที่	ค่า p	ค่า r
1	0.494	0.225	31	0.559	0.282	61	0.656	0.580
2	0.676	0.377	32	0.345	0.204	62	0.664	0.463
3	0.573	0.347	33	0.529	0.550	63	0.344	0.550
4	0.780	0.206	34	0.395	0.558	64	0.635	0.538
5	0.705	0.271	35	0.634	0.505	65	0.440	0.587
6	0.722	0.300	36	0.319	0.336	66	0.743	0.427
7	0.266	0.215	37	0.298	0.227	67	0.672	0.491
8	0.730	0.319	38	0.496	0.383	68	0.498	0.471
9	0.523	0.380	39	0.256	0.375	69	0.411	0.373
10	0.556	0.228	40	0.357	0.445	70	0.328	0.478
11	0.253	0.214	41	0.387	0.271	71	0.510	0.411
12	0.307	0.239	42	0.508	0.546	72	0.261	0.316
13	0.286	0.220	43	0.805	0.503	73	0.373	0.495
14	0.668	0.331	44	0.517	0.333	74	0.295	0.474
15	0.784	0.414	45	0.739	0.232	75	0.286	0.336
16	0.751	0.532	46	0.441	0.298	76	0.282	0.394
17	0.485	0.450	47	0.777	0.398	77	0.564	0.557
18	0.747	0.333	48	0.256	0.440	78	0.357	0.481
19	0.448	0.427	49	0.710	0.419	79	0.324	0.401
20	0.515	0.298	50	0.655	0.486	80	0.286	0.322
21	0.290	0.326	51	0.349	0.482	81	0.353	0.383
22	0.647	0.458	52	0.282	0.306	82	0.610	0.408
23	0.340	0.364	53	0.248	0.252	83	0.573	0.462
24	0.344	0.304	54	0.542	0.552	84	0.461	0.345
25	0.286	0.306	55	0.710	0.430	85	0.249	0.245
26	0.257	0.344	56	0.328	0.404	86	0.527	0.473
27	0.627	0.423	57	0.324	0.393	87	0.556	0.444
28	0.523	0.411	58	0.462	0.499	88	0.419	0.351
29	0.216	0.204	59	0.534	0.448	89	0.290	0.258
30	0.320	0.259	60	0.328	0.240	90	0.324	0.260

ภาคผนวก ค.

ตารางแสดงจำนวนกลุ่มอ้างอิง กลุ่มเปรียบเทียบ ของแบบทดสอบทดสอบวัดความถนัดทางการเรียน
ตามความรู้สึกรู้สึกคุ้นเคย ความรู้สึกพอใจ และความรู้สึกสนใจในข้อสอบ

ตาราง 12 จำนวนกลุ่มอ้างอิง กลุ่มเปรียบเทียบ ของแบบทดสอบวัดความถนัดทางการเรียน

จำนวนข้อ	ความรู้สีก้นเคย			ความรู้สีกสนใจ			ความรู้สีกพอใจ		
	แบบทดสอบวัดความถนัดทางการเรียนด้าน								
	เหตุผล	ภาษา	ตัวเลข	เหตุผล	ภาษา	ตัวเลข	เหตุผล	ภาษา	ตัวเลข
0 ข้อ	77	112	135	13	11	21	7	11	14
1 ข้อ	9	7	8	0	0	1	0	0	3
2 ข้อ	3	4	5	0	2	1	1	1	1
3 ข้อ	8	8	2	3	1	1	2	0	4
4 ข้อ	10	6	5	0	2	0	1	2	0
5 ข้อ	6	6	4	1	0	2	0	2	0
6 ข้อ	8	5	6	2	1	1	2	1	3
7 ข้อ	13	11	7	3	3	6	2	2	3
8 ข้อ	21	9	4	3	0	3	2	1	4
9 ข้อ	22	9	7	1	4	1	7	5	3
10 ข้อ	14	12	14	10	2	1	6	1	5
11 ข้อ	22	7	14	6	8	4	6	5	6
12 ข้อ	22	18	19	14	11	7	9	10	7
13 ข้อ	27	18	5	14	15	5	20	5	3
14 ข้อ	37	25	18	19	12	10	19	18	8
15 ข้อ	28	29	29	27	22	9	19	27	27
16 ข้อ	27	28	25	29	18	21	23	20	15
17 ข้อ	27	22	14	30	25	16	30	25	20
18 ข้อ	26	26	18	26	19	13	20	22	21
19 ข้อ	21	22	12	22	15	16	28	22	17
20 ข้อ	16	22	9	18	18	13	21	16	9
21 ข้อ	20	19	11	13	23	6	18	22	14
22 ข้อ	11	21	12	9	20	9	16	18	13
23 ข้อ	18	10	15	12	12	19	20	21	12
24 ข้อ	13	15	22	18	23	14	19	14	14
25 ข้อ	11	14	9	13	11	16	13	17	21
26 ข้อ	14	12	9	19	16	10	18	9	11
27 ข้อ	9	10	5	21	12	8	21	15	9
28 ข้อ	12	7	11	19	16	17	22	13	11
29 ข้อ	6	8	12	28	22	17	36	24	23
30 ข้อ	27	63	119	192	241	317	177	286	284

กลุ่มอ้างอิง

ภาคผนวก ง.

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจากแบบทดสอบวัดความถนัดทางการเรียน

ตาราง 13 การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผล

ข้อ	เพศ (310 : 285)		ความคุ้นเคย (258 : 307)		ความสนใจ (469 : 116)		ความพอใจ (482 : 103)	
	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA
		α_{MH} ค่า t						
1	1.10	-1.034	0.98	-0.547	0.87	0.103	1.10	0.220
2	0.90	-2.014*	0.98	-0.0358	1.03	0.663	0.69	-1.328
3	0.83	2.850*	1.52*	2.206*	1.58	2.772*	0.82	-0.478
4	1.06	-2.164*	1.41	1.476	0.71	-0.074	1.42	1.455
5	0.96	-1.906	1.37	1.293	0.76	-0.469	1.08	0.635
6	1.15	-1.404	1.02	-0.344	0.73	0.157	0.60	-1.619
7	2.71*	2.394*	1.29	1.042	0.72	-0.175	0.76	-0.562
8	1.65	-0.985	0.68	-1.406	0.73	0.266	0.84	-0.224
9	1.53	-1.626	0.75	-0.934	0.86	0.803	0.91	0.029
10	0.60*	-4.761*	0.97	-0.372	1.18	1.973*	1.32	1.237
11	0.72	-3.857*	0.60	-1.451	1.84	2.546*	1.08	0.134
12	0.90	-3.152*	0.88	-0.718	0.76	0.386	0.66	-0.858
13	0.85	-2.888*	1.05	0.442	1.25	2.055*	1.08	0.514
14	1.27	-1.003	0.76	-1.340	0.95	0.888	0.90	-0.122
15	1.34	0.071	0.71	-2.366*	0.77	-0.114	1.02	0.292
16	1.23	-1.686	0.97	-0.159	0.87	0.970	1.11	0.842
17	0.85	-2.988*	0.85	-0.841	1.49	2.729*	1.77*	2.426*
18	1.60*	1.318	1.04	0.512	1.02	0.385	0.66	-1.628
19	1.72*	1.465	0.70	-2.446*	0.76	-0.386	1.03	0.378
20	0.59*	-5.225*	1.07	0.170	1.24	2.253*	1.59	1.904*
21	0.77	-2.260*	1.18	0.742	1.03	0.716	0.96	-0.044
22	1.10	-2.026*	1.03	0.231	1.17	2.041*	1.06	0.611
23	1.03	-2.530*	1.24	0.927	0.72	0.422	0.90	0.029
24	0.85	-3.521*	1.08	0.735	0.76	0.429	0.98	0.433
25	1.03	-1.454	0.78	-1.987*	1.40	1.917	0.64	-1.126
26	0.91	-2.422*	1.39	1.116	1.07	0.898	2.02*	2.731*
27	0.60*	-4.086*	1.00	0.632	1.09	1.127	0.92	-0.087
28	0.98	-0.629	1.01	0.107	1.07	0.668	1.18	0.795
29	0.47*	-4.743*	0.80	-1.978*	1.24	1.681	0.72	-0.666
30	0.73	-1.753	1.36	1.038	1.76	2.574*	1.72	2.163*

* ข้อสอบที่ทำหน้าที่ต่างกัน
(จำนวนกลุ่มอ้างอิง : กลุ่มเปรียบเทียบ)

ตาราง 14 การตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียน ด้านภาษา

ข้อ	เพศ (310 : 285)		ความคุ้นเคย (299 : 286)		ความสนใจ (491 : 94)		ความพอใจ (494 : 91)	
	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA	วิธี MH	วิธี ANOVA
	α_{MH}	ค่า t	α_{MH}	ค่า t	α_{MH}	ค่า t	α_{MH}	ค่า t
31	1.01	-1.699	1.01	0.375	0.84	0.071	0.78	-0.328
32	0.88	-1.780	1.22	0.692	1.14	0.946	1.29	1.847
33	1.32	-1.140	0.91	-0.659	0.51*	-1.068	0.69	-0.299
34	1.63*	0.145	0.77	-1.389	1.08	0.674	0.98	0.613
35	1.62*	0.338	1.04	0.270	1.07	0.780	0.88	0.372
36	1.29	0.059	1.19	0.285	1.41	1.796	1.04	0.727
37	0.89	-1.794	0.88	-0.459	0.91	0.089	0.71	-0.877
38	0.98	-1.067	0.86	-0.797	1.17	1.135	1.15	1.029
39	0.97	-1.972*	0.74	-0.813	0.70	-0.632	0.98	-0.390
40	1.54*	-0.110	0.91	-0.807	1.02	0.711	0.83	-0.270
41	1.25	0.000	1.32	1.589	0.86	-0.511	0.91	-0.139
42	1.79*	0.054	0.84	-0.698	0.86	0.359	0.59*	-1.115
43	1.01	0.157	1.10	0.667	0.58*	-2.063*	1.07	-0.059
44	0.60*	-4.485*	0.91	-0.573	0.81	-0.367	0.97	0.139
45	0.63*	-3.874*	0.87	-0.573	0.99	0.654	0.90	-0.315
46	0.88	-1.513	0.90	-0.720	1.14	0.864	1.24	0.889
47	0.67	-4.087*	1.31	1.071	0.79	0.913	1.05	1.100
48	1.25	-0.539	1.42	1.169	1.16	0.6+56	1.05	0.653
49	1.15	-0.834	1.10	0.844	1.31	1.470	1.06	0.616
50	1.08	-2.484*	0.80	-0.838	1.27	1.647	1.13	1.215
51	0.71	-3.366*	0.84	-0.760	1.54	1.604	1.07	0.595
52	0.63*	-3.579*	1.00	-0.106	0.83	-0.152	0.96	0.097
53	1.05	-0.649	0.76	-1.345	1.23	0.783	1.06	1.125
54	0.62*	-4.438*	1.18	0.417	1.21	1.507	1.51	2.008*
55	0.58*	-4.889*	0.97	-0.177	1.94*	2.904*	1.37	1.577
56	1.24	-0.593	1.05	0.274	0.97	0.398	0.86	-0.330
57	0.92	-1.897	1.06	0.178	0.77	-0.345	0.72	-0.146
58	1.47	0.864	1.49	1.733	1.36	0.614	1.11	0.744
59	0.76	-3.012*	1.00	-0.335	1.27	1.341	2.36*	3.301*
60	1.16	-0.438	1.19	0.916	0.69	-1.310	0.98	0.544

* ข้อสอบที่ทำหน้าที่ต่างกัน
(จำนวนกลุ่มอ้างอิง : กลุ่มเปรียบเทียบ)

ប្រវត្តិយុវជន

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นางสาวชุตีมา แสงदारารัตน์
วันเดือนปีเกิด	26 พฤษภาคม 2515
สถานที่เกิด	เขตบางกอกน้อย กรุงเทพมหานคร
สถานที่อยู่ปัจจุบัน	107/2 หมู่7 ถนนสุขสวัสดิ์ แขวงราษฎร์บูรณะ เขตราษฎร์บูรณะ กรุงเทพมหานคร 10140
ตำแหน่งหน้าที่การงาน	หัวหน้าแผนกทะเบียนและประมวลผล
สถานที่ทำงานปัจจุบัน	วิทยาลัยเซนต์หลุยส์ 215/4 ถนนสาทรใต้ ยานนาวา สาทร กรุงเทพมหานคร 10210
ประวัติการศึกษา	
พ.ศ. 2531	มัธยมศึกษาตอนปลาย จากโรงเรียนสตรีวิทยา
พ.ศ. 2535	บธ.บ (บริหารงานอุตสาหกรรม) จากมหาวิทยาลัยหอการค้าไทย
พ.ศ. 2546	กศ.ม (การวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ