

การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา
ตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน

ปริญญานิพนธ์

ของ

วิภา จำนุ่น

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

กันยายน 2544

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา
ตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน

บทคัดย่อ

ของ

วิภา จ๋ามั่น

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

กันยายน 2544

วิภา จำนั. (2544). การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา ตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์อังคณา สายยศ, รองศาสตราจารย์ชูศรี วงศ์รัตน์.

การวิจัยครั้งนี้ มีจุดมุ่งหมายเพื่อเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา เมื่อตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซลในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง โดยเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงและค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ในการศึกษาครั้งนี้ มีกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 1,041 คน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น ผลการวิจัยมีดังนี้

1. จำนวนข้อสอบที่มีความลำเอียง เมื่อตรวจสอบด้วยวิธีชิปเทสท์ ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนข้อสอบที่มีความลำเอียง ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และระหว่างกลุ่มข้อสอบที่มีระดับความง่ายปานกลางกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อตรวจสอบด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า ข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. จำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง เมื่อตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

A COMPARISON OF ITEM BIAS DETECTION OF VERBAL ABILITY TEST
USING DIFFERENT ITEM EASINESS AND METHODS

AN ABSTRACT

BY

WIPA JUMMUN

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

September 2001

Wipa Jummun. (2001). *A comparison of item bias detection of verbal ability test using different item easiness and methods*. Master Thesis, M.Ed. (Educational Measurement). Bangkok : Graduate School Srinakharinwirot University. Advisor Committee : Assoc. Prof. Aungkana Saiyos, Assoc. Prof. Chusri Wongrattana.

The purpose of this research was to compare the item bias detection of verbal ability test through low, medium and high item easiness by using SIBTEST method and Mantel-Haenszel method. The comparison of items bias were to determine the number of biased items, the reliability of the test after deleted biased items. The sample group of this study was 1,041 mathayomsuksa 3 students selected by stratified random sampling. The result of the study were as following :

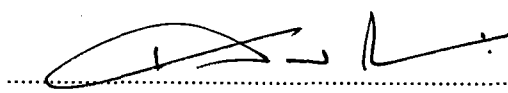
1. The number of biased items detected by SIBTEST method between low and high item easiness were statistical significant at .05 level but low and medium, medium and high item easiness were not statistical significant. When detected by Mantel-Haenszel method, the number of biased items among three item easiness were not statistical significant.
2. The number of biased items between detected by SIBTEST and Mantel-Haenszel methods with low, medium and high item easiness were not statistical significant
3. The reliability of test with no biased items, between low, medium and high item easiness detected by SIBTEST and Mantel-Haenszel methods were not statistical significant.
4. The reliability of test with no biased items, between detected by SIBTEST and Mantel-Haenszel methods with low, medium and high item easiness were not statistical significant.

ปริญญานิพนธ์
เรื่อง

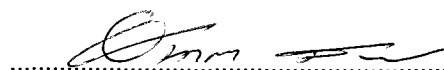
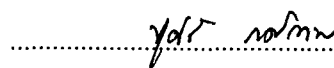
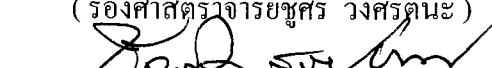
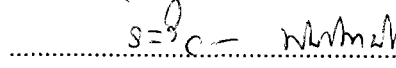
การเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา
ตามระดับความง่ายของข้อสอบและวิธีที่ต่างกัน

ของ
นางสาววิภา จำนัณ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)
วันที่ 27 เดือน กันยายน พ.ศ. 2544

คณะกรรมการสอบปริญญานิพนธ์

..... ประธาน
(รองศาสตราจารย์อังคณา สายยศ)
..... กรรมการ
(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)
..... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ชวลิต รวยอาจิน)
..... กรรมการที่แต่งตั้งเพิ่มเติม
(ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยความอนุเคราะห์อย่างยิ่งจากรองศาสตราจารย์
อังคณา สายยศ ประธานกรรมการควบคุมปริญญานิพนธ์ รองศาสตราจารย์ชูศรี วงศ์รัตนะ
กรรมการควบคุมปริญญานิพนธ์ อาจารย์ชวลิต รวยอาจิน และ ผู้ช่วยศาสตราจารย์ระวีวรรณ
พันธ์พานิช คณะกรรมการแต่งตั้งเพิ่มเติมในการสอบปากเปล่าปริญญานิพนธ์ ที่ได้กรุณาให้
คำแนะนำและข้อคิดเห็นต่าง ๆ อันเป็นประโยชน์อย่างยิ่งในการปรับปรุงแก้ไขให้การวิจัยครั้งนี้
สมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณรองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์, รองศาสตราจารย์
นิภา ศรีไพโรจน์, รองศาสตราจารย์วิญญา วิศาลาภรณ์, ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธ์พานิช
และ อาจารย์ชวลิต รวยอาจิน ที่ให้ความกรุณาเป็นผู้เชี่ยวชาญในการตรวจสอบเครื่องมือวิจัย
ตลอดจนให้คำแนะนำและข้อคิดเห็นต่าง ๆ ในการแก้ไขเครื่องมือวิจัย

ขอกราบขอบพระคุณผู้บริหาร คณะครู อาจารย์ทุกท่านของโรงเรียนที่เป็นกลุ่มตัวอย่าง
และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บข้อมูลเป็นอย่างดี

ขอขอบคุณเพื่อน ๆ ชาววัดผลการศึกษา รวมทั้งกัลยาณมิตรทุกคนที่ให้ความช่วยเหลือ
และเป็นกำลังใจในการทำวิจัยครั้งนี้

กราบขอบพระคุณบิดา มารดา และขอขอบคุณพี่น้อง ตลอดจนผู้ที่มีส่วนเกี่ยวข้องทุกท่าน
ที่เป็นกำลังใจสนับสนุนช่วยเหลือ ทำให้สามารถทำปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชา
พระคุณบิดา มารดา และ ครู อาจารย์ ตลอดจนผู้มีพระคุณที่ได้อบรมสั่งสอนตลอดมา

วิภา จำมัน

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการศึกษาค้นคว้า.....	4
ความสำคัญของการศึกษาค้นคว้า.....	5
ขอบเขตของการศึกษาค้นคว้า.....	5
ประชากรและกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า.....	5
ตัวแปรที่ศึกษา.....	5
นิยามศัพท์เฉพาะ.....	6
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	8
ความหมายของข้อสอบที่ลำเอียง.....	8
การตรวจสอบข้อสอบที่ลำเอียง.....	9
หลักการวิเคราะห์ค่าความง่ายของข้อสอบ.....	16
งานวิจัยที่เกี่ยวข้อง.....	17
สมมติฐานของการศึกษาค้นคว้า.....	27
3 วิธีดำเนินการศึกษาค้นคว้า.....	28
ประชากร.....	28
กลุ่มตัวอย่าง.....	28
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	29
การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	31
ตัวอย่างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	33
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	36
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	36

สารบัญ (ต่อ)

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล.....	41
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	41
การเสนอผลการวิเคราะห์ข้อมูล.....	42
ผลการวิเคราะห์ข้อมูล.....	42
5 สรุป อภิปรายผล และ ข้อเสนอแนะ.....	49
ความมุ่งหมายของการศึกษาค้นคว้า.....	49
กลุ่มตัวอย่าง.....	49
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	49
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	50
การเสนอผลการวิเคราะห์ข้อมูล.....	50
สรุปผลการวิเคราะห์ข้อมูล.....	51
อภิปรายผล.....	51
ข้อเสนอแนะ.....	54
บรรณานุกรม.....	55
ภาคผนวก.....	60
ประวัติย่อผู้วิจัย.....	85

บัญชีตาราง

ตาราง	หน้า
1 ข้อมูลการสอบของกลุ่มตัวอย่างกลุ่ม R และ กลุ่ม F ที่มีคะแนนรวม อยู่ในช่วงคะแนน j ของข้อสอบข้อที่ i.....	14
2 การแปลความหมายของดัชนีค่าความง่าย (p_e) ของข้อสอบ.....	16
3 จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามขนาดโรงเรียน.....	29
4 การตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนก ตามระดับความง่ายของข้อสอบและวิธีตรวจสอบความลำเอียงที่ต่างกัน.....	43
5 การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างระดับความง่ายของข้อสอบ ต่างกัน โดยใช้การตรวจสอบความลำเอียงวิธีเดียวกัน.....	45
6 การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงระหว่างการตรวจสอบความลำเอียง ด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน.....	46
7 ค่าความเชื่อมั่นของแบบทดสอบ ค่า Fisher-Z และค่าการทดสอบนัยสำคัญของค่า ความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างระดับความง่ายของข้อสอบต่างกัน โดยใช้การตรวจสอบความลำเอียง วิธีเดียวกัน.....	47
8 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความ ลำเอียงออกแล้ว ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับ ความง่ายของข้อสอบเหมือนกัน.....	48
9 สรุปจำนวนข้อของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่าย ของข้อสอบ.....	71
10 ค่าความง่าย (p_e) และ ค่าอำนาจจำแนก (r) ของแบบทดสอบวัดความสามารถ ด้านภาษา จำนวน 60 ข้อ.....	73
11 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายต่ำ จำนวน 20 ข้อ จากการ ตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซล...	80

บัญชีตาราง (ต่อ)

ตาราง

หน้า

- 12 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายปานกลาง จำนวน 20 ข้อ
จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์ และวิธี
แมนเทิล-แฮนส์เซล..... 81
- 13 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายสูง จำนวน 20 ข้อ จากการ
ตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์และวิธีแมนเทิล-แฮนส์เซล... 82

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1 ลำดับขั้นในการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	30
---	----

บทที่ 1

บทนำ

ภูมิหลัง

การทดสอบเป็นกระบวนการที่สำคัญอย่างหนึ่งของการวัดผลการศึกษา จนอาจกล่าวได้ว่า ที่ใดมีการสอน ที่นั่นมีการสอบ (ชวาล แพรัตกุล. 2514 : 1) ซึ่งการทดสอบหรือการสอบนี้ จะจัดให้มีขึ้นในทุกๆ ระดับชั้น ตั้งแต่ชั้นอนุบาลจนถึงระดับอุดมศึกษา ไม่ว่าจะเป็นโรงเรียนของภาครัฐหรือภาคเอกชน รวมไปถึงการสอบที่กระทำทั้งในโรงเรียนและนอกโรงเรียนด้วย ซึ่งในการทดสอบนั้นมีปรัชญาที่เกี่ยวกับการทดสอบว่า การทดสอบเป็นการค้นหาและพัฒนาสมรรถภาพของมนุษย์ (ชวาล แพรัตกุล. 2518 : 34 - 129) แสดงว่า การสอบมีความมุ่งหมายอยู่ 2 ประการ คือ เพื่อตรวจสอบความรู้ ความสามารถของผู้เรียน และเพื่อกระตุ้นสมรรถภาพต่าง ๆ ของผู้เรียน หรือเพื่อพัฒนานักเรียนให้เกิดการเรียนรู้ที่สูงขึ้น (พรณี จินตมาศ. 2540 : 1)

เครื่องมือส่วนใหญ่ที่นำมาใช้ในการวัดผลการศึกษาหรือการสอบ ก็คือ แบบทดสอบ เนื่องจากมีความสะดวกในการนำมาใช้ ครูและนักการศึกษาส่วนใหญ่มีความเข้าใจและคุ้นเคยกับเครื่องมือประเภทนี้เป็นอย่างดี รวมถึงขั้นตอนการสร้างก็ไม่ยุ่งยากจนเกินไป หากได้ทำการศึกษา ลักษณะเด่นของแบบทดสอบแต่ละชนิดและวิธีการสร้าง จะทำให้สามารถเลือกใช้ได้เหมาะสมกับจุดมุ่งหมายที่ต้องการวัดได้ (ญานภัทร สีหะมงคล. 2540 : 1) แต่การวัดผลการศึกษาจะมีคุณภาพหรือไม่ หรือสามารถวัดได้ครอบคลุมเนื้อหาเพียงใด ย่อมขึ้นอยู่กับวิธีการวัดและเครื่องมือที่ใช้ในการวัดเป็นสำคัญ ถ้าเครื่องมือที่ใช้มีคุณภาพดี ได้รับการตรวจสอบคุณภาพมาแล้วทั้งในด้านความเที่ยงตรง ความเชื่อมั่น ค่าอำนาจจำแนก รวมทั้งระดับความยากง่ายของข้อสอบ ก็จะทำให้สามารถวัดผลได้ถูกต้อง ตรงตามความเป็นจริงและมีความน่าเชื่อถือ ซึ่งจะส่งผลให้การประเมินผลและการตัดสินใจต่าง ๆ มีความถูกต้องและมีประสิทธิภาพ ดังนั้น ผู้สร้างแบบทดสอบจึงควรคำนึงถึงคุณภาพของแบบทดสอบให้มีคุณภาพที่น่าเชื่อถือได้ มีความคลาดเคลื่อนในการวัดน้อยที่สุด และเกิดความยุติธรรมกับผู้สอบในทุก ๆ กลุ่ม ไม่ให้เกิดการได้เปรียบเสียเปรียบกับกลุ่มใดกลุ่มหนึ่ง โดยเฉพาะ ไม่ว่าผู้สอบจะมีความแตกต่างกันในด้านเพศ ภาษา เชื้อชาติ วัฒนธรรม หรือศาสนา ซึ่งลักษณะเหล่านี้เป็นคุณสมบัติของข้อสอบที่ปราศจากความลำเอียง ดังนั้น ข้อสอบที่ดี จึงควรเป็นทั้งข้อสอบที่มีคุณภาพและข้อสอบที่ไม่มีความลำเอียง

การศึกษาเกี่ยวกับข้อสอบที่มีความลำเอียงนั้น ได้รับการพัฒนามาอย่างต่อเนื่องทั้งในประเทศไทยและในต่างประเทศ สำหรับในประเทศไทยนั้น จากการศึกษาผลงานวิจัยที่ผ่านมา

พบว่า จะทำการศึกษาเปรียบเทียบข้อสอบที่มีความลำเอียงที่เกิดจากความแตกต่างระหว่างกลุ่ม และเปรียบเทียบผลการตรวจสอบความลำเอียงด้วยวิธีที่แตกต่างกัน โดยทำการศึกษาในด้านของ จำนวนข้อสอบที่มีความลำเอียงและเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือก ข้อสอบที่มีความลำเอียงออกแล้ว ดังเช่นงานวิจัยของ พัทรี ปิยภณท์ (2531 : 66) ที่ได้ทำการศึกษา ความลำเอียงของข้อสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ในระดับชั้น ประถมศึกษาปีที่ 6 ระหว่างกลุ่มนักเรียนที่อยู่ในสภาพแวดล้อมต่างกัน คือ นักเรียนใน กรุงเทพมหานครและนักเรียนที่อยู่ในจังหวัดสมุทรสาคร พบว่า ข้อสอบมีความลำเอียงถึง 18 ข้อ จากข้อสอบจำนวน 45 ข้อ และจากงานวิจัยของ คมศักดิ์ ชื่นชม (2539 : 108) ที่ได้ทำการศึกษา เปรียบเทียบผลการวิเคราะห์ความลำเอียงของแบบทดสอบวัดจริยธรรมความซื่อสัตย์ ตามทฤษฎี คลาสสิกคอลที่ใช้วิธีวิเคราะห์ต่างกัน 3 วิธี คือ วิธีค่าอำนาจจำแนก วิธีวิเคราะห์องค์ประกอบ และ วิธีวิเคราะห์การถดถอย ซึ่งผลการวิจัย พบว่า วิธีวิเคราะห์ทั้ง 3 วิธี พบจำนวนข้อสอบที่มีความ ลำเอียงน้อยและไม่มีความแตกต่างกันที่ระดับนัยสำคัญ .05 นอกจากนี้ ยังมีผลงานวิจัยอีกหลายเรื่อง ที่ได้ศึกษาเกี่ยวกับข้อสอบที่ลำเอียง ได้แก่ เจน จันทรแสง. 2537 ; นิรมล ชัยชวลิต. 2537 ; จิตสุดา ธารพร. 2539 ; อังคณา สายยศ. 2539 ; พรรณี จินมาศ. 2540 ; นวลอนงค์ นันทวงศ์. 2543 เป็นต้น

สำหรับการศึกษาเรื่องข้อสอบที่ลำเอียงในต่างประเทศนั้น ส่วนใหญ่แล้วจะทำการศึกษา เกี่ยวกับเทคนิควิธีการตรวจสอบความลำเอียงเพื่อเปรียบเทียบค่าพารามิเตอร์ของข้อสอบระหว่าง สองกลุ่มย่อยว่ามีความแตกต่างกันหรือไม่ และศึกษาถึงปัจจัยที่ส่งผลต่อการตรวจสอบความลำเอียง ของข้อสอบด้วย (อารี วัชรโสติกุล. 2543 : 3) เช่น ขนาดของกลุ่มตัวอย่าง ความยาวของแบบ ทดสอบ ระดับความยาก และค่าอำนาจจำแนกของแบบทดสอบ รวมไปถึงการเปรียบเทียบการ ตรวจสอบความลำเอียงด้วยวิธีที่แตกต่างกันด้วย ดังเช่นงานวิจัยของ เชพเพิร์ด และคณะ (Berk. 1982 : 169 ; citing Shepard Camilli & Averrill. 1981. *Journal of Educational Statistics*.) ที่ได้ทำ การศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียง 5 วิธี คือ วิธีไคสแควร์ วิธีแปลงค่าความยาก วิธี พอยท์ไบซีเรียล วิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีไค้่งลักษณะข้อสอบ 1 พารามิเตอร์ ผลการวิจัยพบว่า วิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ เป็นวิธีที่เหมาะสมที่สุดในทางทฤษฎี มี ความถูกต้องและตรงที่สุด วิธี ไคสแควร์เป็นวิธีที่มีความสอดคล้องกับวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ ในทางปฏิบัติสามารถนำมาใช้แทนวิธีไค้่งลักษณะข้อสอบ 3 พารามิเตอร์ได้ นอกจากนี้ ยังมีงานวิจัยของ เมเซอร์ คลอเซอร์ และแฮมเบิลตัน (Mazor , Clauser & Hambleton. 1992 : 443 - 451) ที่ได้ศึกษาผลกระทบของขนาดของกลุ่มตัวอย่างที่มีต่อการตรวจสอบความลำเอียงของข้อสอบ

ด้วยวิธีแมนเทล-แฮนส์เซล โดยใช้ข้อมูลที่จำลองด้วยโปรแกรม DATAGEN แบบ 3 พารามิเตอร์ โดยขนาดกลุ่มตัวอย่างที่ศึกษามี 5 ขนาด คือ 100, 200, 500, 1,000 และ 2,000 คน ผลการศึกษาพบว่า เมื่อใช้กลุ่มตัวอย่างขนาด 2,000 คน การตรวจสอบความลำเอียงผิดพลาดร้อยละ 50 แต่เมื่อใช้จำนวนผู้สอบน้อยกว่า 500 คน จะตรวจสอบความลำเอียงได้ผิดพลาดร้อยละ 30 ส่วนข้อสอบที่ยากมาก ข้อสอบที่มีค่าอำนาจจำแนกต่ำ และข้อสอบที่มีค่าความยากต่างกันเพียงเล็กน้อยสำหรับสองกลุ่ม จะไม่สามารถตรวจสอบความลำเอียงได้

สำหรับวิธีที่ใช้ในการตรวจสอบข้อสอบที่ลำเอียงนั้นมีอยู่หลายวิธีด้วยกัน โดยส่วนใหญ่แล้ว จะยึดตามแนวทฤษฎี 2 ทฤษฎีเป็นหลัก ได้แก่ ทฤษฎีการวัดผลแบบดั้งเดิม (Classical Test Theory : CTT) และทฤษฎีการตอบข้อคำถาม (Item Response Theory : IRT) สำหรับวิธีที่ยึดหลักของทฤษฎีการวัดผลแบบดั้งเดิม เช่น วิธีแปลงค่าความยาก (Item Transformed Difficulty) วิธีวิเคราะห์ความแปรปรวน (Analysis of Variance) วิธีวิเคราะห์การถดถอย (Regression Analysis) การวิเคราะห์ไคสแควร์ (Chi-Square) วิธีวิเคราะห์องค์ประกอบ (Factor Analysis) วิธีล็อกลิเนียร์ (Loglinear) และวิธีการถดถอยโลจิสติก (Logistic Regression) ส่วนวิธีที่ยึดหลักของทฤษฎีการตอบข้อคำถาม เช่น วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ (Item Characteristic Curve - 3 Parameter) วิธีโค้งลักษณะข้อสอบ 2 พารามิเตอร์ (Item Characteristic Curve - 2 Parameter) และ วิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ (Item Characteristic Curve - 1 Parameter) โดยแต่ละวิธีตรวจสอบความลำเอียงนี้ต่างก็มีทั้งข้อดีและข้อเสียที่แตกต่างกัน แต่จากผลการวิจัยส่วนใหญ่ พบว่า การตรวจสอบข้อสอบที่ลำเอียงโดยใช้หลักของทฤษฎีการตอบข้อคำถามนั้น เป็นวิธีที่ให้ผลการวิเคราะห์ที่ถูกต้องที่สุด แต่การใช้ทฤษฎีการตอบข้อคำถามในการตรวจสอบข้อสอบที่ลำเอียงนั้น เป็นวิธีที่มีการคำนวณที่ยากและซับซ้อน ใช้กลุ่มตัวอย่างขนาดใหญ่ และสิ้นเปลืองค่าใช้จ่ายมาก ทำให้เกิดความยุ่งยากในทางปฏิบัติ จากเหตุผลดังกล่าว จึงมีผู้เสนอวิธีการตรวจสอบความลำเอียงโดยใช้สถิติแบบนอนพารามेटริกในการตรวจสอบข้อสอบที่ลำเอียงขึ้น เพื่อหลีกเลี่ยงการใช้ทฤษฎีการตอบข้อคำถาม นั่นคือ วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel) และวิธีซิปเทสต์ (SIBTEST) ซึ่งวิธีการทั้งสองนี้มีวิธีการคำนวณที่ไม่ยุ่งยากซับซ้อนและสิ้นเปลืองค่าใช้จ่ายไม่มาก เหมาะที่จะนำมาใช้ในทางปฏิบัติ (พรณี จินตมาศ. 2540 : 3) และนอกจากนี้ ยังสามารถวิเคราะห์ได้ด้วยคอมพิวเตอร์ส่วนบุคคล จึงเป็นที่สนใจและนิยมใช้ในปัจจุบัน

จากข้อความที่กล่าวมาแล้วข้างต้น จะเห็นว่า กลุ่มผู้สอบที่มีความแตกต่างกันในด้านต่าง ๆ ไม่ว่าจะเป็นความแตกต่างทางด้านเชื้อชาติ ภาษา ศาสนา และวัฒนธรรม รวมทั้งสภาพภูมิศาสตร์ และสภาพสังคม เป็นสาเหตุที่ทำให้ข้อคำถามบางข้อเกิดความลำเอียงกับผู้สอบกลุ่มใดกลุ่มหนึ่ง โดยเฉพาะ ส่วนวิธีที่ใช้ในการตรวจสอบข้อสอบที่ลำเอียงนั้นก็มีอยู่หลายวิธีด้วยกัน ซึ่งงานวิจัย

ส่วนใหญ่จะทำการศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงด้วยวิธีต่าง ๆ ที่แตกต่างกัน ด้วยเหตุผลดังกล่าวนี้ ผู้วิจัยจึงมีความสนใจที่จะศึกษาการเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา ที่มีระดับความง่ายของข้อสอบแตกต่างกัน โดยแบ่งเป็นข้อสอบที่มีระดับความง่ายต่ำ ข้อสอบที่มีระดับความง่ายปานกลาง และข้อสอบที่มีระดับความง่ายสูง ซึ่งใช้วิธีชิปเทสต์ และวิธีแมนเทล-เฮนส์เซลในการตรวจสอบข้อสอบที่ลำเอียง โดยศึกษาในด้านจำนวนข้อสอบที่มีความลำเอียง และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว อันจะเป็นแนวทางในการพัฒนาและปรับปรุงในการสร้างแบบทดสอบให้มีคุณภาพต่อไป

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน
2. เพื่อเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน
3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน
4. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาครั้งนี้ ทำให้ทราบว่า การตรวจสอบความลำเอียงด้วยวิธีใดที่จะพบจำนวนข้อสอบที่มีความลำเอียงต่อเพศได้มากที่สุด และข้อคำถามของแบบทดสอบจะเกิดความลำเอียงเมื่อมีค่าความง่ายของข้อสอบอยู่ในระดับใด รวมทั้งทราบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ซึ่งจะเป็นแนวทางในการตัดสินใจพัฒนาการสร้างแบบทดสอบที่มีคุณภาพและมีความยุติธรรมต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 25 โรงเรียน 133 ห้องเรียน เป็นนักเรียนชาย 2,574 คน เป็นนักเรียนหญิง 2,687 คน รวมทั้งหมด 5,261 คน

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) และห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้จำนวนห้องเรียน 26 ห้องเรียน ที่มีจำนวนนักเรียนทั้งสิ้น 1,041 คน

3. ตัวแปรที่ศึกษา

1) ตัวแปรอิสระ มีดังนี้

1.1 วิธีตรวจสอบข้อสอบที่ลำเอียง ได้แก่

1.1.1 วิธีชิปเทสต์ (SIBTEST)

1.1.2 วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel)

1.2 ระดับความง่ายของข้อสอบ ได้แก่

1.2.1 ข้อสอบที่มีระดับความง่ายต่ำ

1.2.2 ข้อสอบที่มีระดับความง่ายปานกลาง

1.2.3 ข้อสอบที่มีระดับความง่ายสูง

2) ตัวแปรตาม ได้แก่ ผลการตรวจสอบข้อสอบที่ลำเอียง จำแนกเป็น

2.1 จำนวนข้อสอบที่ลำเอียง

2.2 ความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียง

ออกแล้ว

3) ตัวแปรที่ใช้ในการตรวจสอบข้อสอบที่ลำเอียง ได้แก่ เพศชาย และเพศหญิง

นิยามศัพท์เฉพาะ

1. ข้อสอบที่ลำเอียง หมายถึง ข้อสอบที่วัดความสามารถหรือคุณลักษณะเดียวกัน แต่เมื่อนำไปสอบกับผู้สอบที่มีคุณลักษณะแตกต่างกัน 2 กลุ่ม ที่มีความสามารถเท่ากันแล้ว ได้ผลการสอบแตกต่างกัน ซึ่งในที่นี้ คือ กลุ่มผู้สอบเพศชาย และกลุ่มผู้สอบเพศหญิง

2. การตรวจสอบข้อสอบที่ลำเอียง หมายถึง วิธีการค้นหาข้อสอบที่มีความลำเอียงซึ่งในการวิจัยครั้งนี้ ใช้ 2 วิธี ดังนี้

2.1 วิธีซิปเทสท์ (SIBTEST) หมายถึง วิธีการวิเคราะห์หาดัชนีความลำเอียงของข้อสอบโดยพิจารณาค่าความลำเอียง β ที่ได้จากการทดสอบความแตกต่างของการตอบถูกของผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ โดยข้อสอบที่มีความลำเอียงจะมีค่า B_U หรือ Z_{SIB} แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2.2 วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel) หมายถึง วิธีการวิเคราะห์หาดัชนีความลำเอียงของข้อสอบโดยพิจารณาเปรียบเทียบสัดส่วนการตอบข้อสอบของกลุ่มผู้สอบ 2 กลุ่ม คือ กลุ่มอ้างอิงและกลุ่มเปรียบเทียบ โดยข้อสอบที่มีความลำเอียงจะมีค่า α_{MH} ไม่เท่ากับ 1 และค่า χ^2_{MH} มากกว่าค่า χ^2 ที่ $df = 1$ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. กลุ่มอ้างอิง (Reference Group : R) หมายถึง กลุ่มผู้สอบที่คาดว่าจะจะเป็นกลุ่มที่ได้เปรียบจากข้อสอบที่มีความลำเอียง นั่นคือ เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ลำเอียงได้ถูกต้องมากกว่าผู้สอบอีกกลุ่มหนึ่งที่ได้คะแนนจากแบบทดสอบเท่ากัน ซึ่งกลุ่มอ้างอิงในการวิจัยครั้งนี้ คือ กลุ่มผู้สอบเพศชาย

4. กลุ่มเปรียบเทียบ(Focal Group : F) หมายถึง กลุ่มผู้สอบที่คาดว่า จะเป็นกลุ่มที่เสียเปรียบจากข้อสอบที่มีความลำเอียง นั่นคือ เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ลำเอียงได้ถูกต้องต่ำกว่าผู้สอบกลุ่มอ้างอิงที่ได้คะแนนจากแบบทดสอบเท่ากัน ซึ่งกลุ่มเปรียบเทียบในการวิจัยครั้งนี้ คือ กลุ่มผู้สอบเพศหญิง

5. ผลการตรวจสอบข้อสอบที่ลำเอียง หมายถึง ผลที่ได้จากการตรวจสอบข้อสอบที่ลำเอียง ด้วยวิธีการตรวจสอบแต่ละวิธี (ตามข้อ 2) จำแนกได้ดังนี้

5.1 จำนวนข้อสอบที่ลำเอียง หมายถึง จำนวนข้อสอบที่ให้ผลการสอบที่แตกต่างกัน ระหว่างกลุ่มผู้สอบเพศชายและกลุ่มผู้สอบเพศหญิง โดยวัดได้จากแบบทดสอบที่ผู้วิจัยสร้างขึ้น

5.2 ค่าความเชื่อมั่น (Reliability) ของแบบทดสอบ หมายถึง คุณสมบัติของแบบทดสอบที่วัดความรู้ ความสามารถของผู้สอบได้อย่างคงที่แน่นอน ซึ่งในการวิจัยครั้งนี้จะศึกษาค่าความเชื่อมั่นของแบบทดสอบหลังการคัดข้อสอบที่ลำเอียงออกแล้ว โดยคำนวณจากสูตรของ กูเดอร์ ริชาร์ดสัน (KR-20)

6. ระดับความง่ายของข้อสอบ หมายถึง สัดส่วนของผู้ที่ตอบถูกในข้อสอบข้อหนึ่ง ๆ ซึ่งแทนด้วยดัชนีค่าความง่าย (p_e) โดยแบ่งเป็น 3 ระดับ คือ

6.1 ระดับความง่ายต่ำ หมายถึง ข้อสอบที่มีดัชนีค่าความง่าย (p_e) ตั้งแต่ 0.20 – 0.39

6.2 ระดับความง่ายปานกลาง หมายถึง ข้อสอบที่มีดัชนีค่าความง่าย (p_e) ตั้งแต่ 0.40 – 0.59

6.3 ระดับความง่ายสูง หมายถึง ข้อสอบที่มีดัชนีค่าความง่าย (p_e) ตั้งแต่ 0.60 – 0.80

7. แบบทดสอบวัดความสามารถด้านภาษา หมายถึง แบบทดสอบที่วัดความสามารถเกี่ยวกับภาษา โดยจำแนกเป็น 5 ด้าน ดังนี้

7.1 ด้านคำตรงข้าม หมายถึง แบบทดสอบที่วัดความสามารถในการแปลความหมายของคำหรือข้อความในทิศทางตรงกันข้ามกับคำหรือข้อความที่กำหนดให้

7.2 ด้านคำที่มีความหมายใกล้เคียง หมายถึง แบบทดสอบที่วัดความสามารถในการเลือกหาความหมายของคำ โดยการเปรียบเทียบและวิเคราะห์ดูว่า ในชุดของคำ หรือข้อความที่มีอยู่นั้น คำหรือข้อความใดมีความหมายใกล้เคียงกับคำหรือข้อความที่กำหนดให้มากที่สุด

7.3 ด้านคำศัพท์สัมพันธ์ หมายถึง แบบทดสอบที่วัดความสามารถในการค้นหาคำที่เกี่ยวข้องระหว่างคำต่าง ๆ ว่าคำใดมีความเกี่ยวข้องกับคำที่กำหนดให้มากที่สุด

7.4 ด้านความหมายไม่เข้าพวก หมายถึง แบบทดสอบที่วัดความสามารถในการพิจารณาว่าประโยคหรือวลีต่าง ๆ ที่กำหนดให้ในข้อใดที่มีความหมายแตกต่างหรือผิดแปลกออกไปจากข้ออื่น ๆ

7.5 ด้านความเข้าใจภาษา หมายถึง แบบทดสอบที่วัดความสามารถในการอ่านข้อความเรื่องราว หรือสถานการณ์ต่าง ๆ แล้วสามารถจับใจความสำคัญ แปลความ ตีความ และขยายความจากข้อความ เรื่องราว หรือสถานการณ์ต่าง ๆ ที่กำหนดให้ได้อย่างรวดเร็วและถูกต้อง

8. ผู้เชี่ยวชาญ หมายถึง นักวัดผลการศึกษาที่มีประสบการณ์ในการสอนมาแล้ว อย่างน้อย 5 ปี จำนวน 5 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า ผู้วิจัยได้ศึกษาเนื้อหาต่าง ๆ ดังนี้

1. ความหมายของข้อสอบที่ลำเอียง
2. การตรวจสอบข้อสอบที่ลำเอียง
3. หลักการวิเคราะห์ค่าความง่ายของข้อสอบ
4. งานวิจัยที่เกี่ยวข้อง

1. ความหมายของข้อสอบที่ลำเอียง

คำว่า ข้อสอบที่ลำเอียง (Item Bias) หรือบางงานวิจัยอาจใช้คำว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Function : DIF) ได้มีผู้ให้ความหมายไว้หลายท่านด้วยกัน ดังนี้

อลเลน และเยน (Allen and Yen. 1979 : 129) ได้ให้ความหมายว่า ข้อสอบที่ลำเอียงเป็นข้อสอบที่วัดความสามารถหรือคุณลักษณะทางจิตวิทยาของผู้สอบ ที่มีระดับความสามารถเท่ากันแล้วได้ผลต่างกัน โดยพิจารณาจากโครงสร้างข้อสอบของผู้สอบในแต่ละกลุ่มจะต่างกัน เช่นเดียวกับแฮมเบิลตัน และคณะ (Hambleton, et al. 1991 : 110) แมคคอลเลย์ และเมนโดซา (McCauley and Mendoza. 1985 : 389) อังคณา สายยศ (2539 : 6) ที่ให้ความหมายของข้อสอบที่ลำเอียงว่า พิจารณาจาก ผลการสอบที่แตกต่างกันจากผู้สอบสองกลุ่มที่มีความสามารถเท่ากัน

เอลส์ และคณะ (Jensen. 1980 : 369 ; citing Eells, et al. 1951. *Intelligence and cultural differences.*) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงทางวัฒนธรรมว่า หมายถึงความแตกต่างกันในขอบเขตของการได้รู้จักและคุ้นเคยกับเนื้อหาวิชาเฉพาะหรือกระบวนการเฉพาะ ซึ่งผู้ที่ถูกทดสอบนั้น หากมีกลุ่มใดได้รู้จักหรือคุ้นเคยกับข้อสอบนั้นมาก่อน ย่อมมีโอกาสตอบถูกมากกว่าผู้สอบคนอื่นที่มีลักษณะของกลุ่มแตกต่างกัน

คามิลลี และเชพเพิร์ด (Camilli and Shepard. 1994 : 8) กล่าวว่า ข้อสอบที่ลำเอียงเป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Error) ชนิดหนึ่งของการวัดที่ทำให้เกิดการได้เปรียบเสียเปรียบต่อสมาชิกกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ เช่น เอื้อต่อเพศชายมากกว่าเพศหญิง เอื้อต่อกลุ่มคนผิวขาวมากกว่าผิวดำ เป็นต้น

มิลล์แซป และอีเวอร์สัน (Millsap and Everson. 1993 : 297) กล่าวว่า ข้อสอบที่ลำเอียงเกิดจากระบบของการวัดที่ไม่แม่นยำ (Inaccuracy)

จากความหมายดังกล่าว สรุปได้ว่า ข้อสอบที่ลำเอียง หมายถึง คุณลักษณะของข้อสอบที่เมื่อนำไปสอบกับกลุ่มตัวอย่างที่มีความสามารถเท่าเทียมกันแล้ว ได้ผลการสอบที่แตกต่างกัน อันเนื่องมาจากความแตกต่างในด้านใดด้านหนึ่งของกลุ่มผู้สอบแต่ละกลุ่ม เช่น เพศ ศาสนา เชื้อชาติ สภาพภูมิศาสตร์ เป็นต้น

2. การตรวจสอบข้อสอบที่ลำเอียง

การตรวจสอบข้อสอบที่ลำเอียงจะกระทำในสองขั้นตอน คือ ขั้นตอนแรกเป็นขั้นตอนของการสร้างแบบทดสอบ โดยนำแบบทดสอบที่สร้างขึ้นไปให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงของแบบทดสอบว่าวัดคุณลักษณะเดียวกันหรือไม่ทั้งในด้านรูปแบบ เนื้อหา ภาษาที่ใช้ในข้อคำถามและตัวเลือกต่างๆ ต้องมีความสอดคล้องกัน เพื่อไม่ให้เกิดการได้ประโยชน์หรือเสียประโยชน์ต่อกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ ส่วนขั้นตอนต่อมา จะเป็นการตรวจสอบความลำเอียงของข้อสอบหลังจากที่นำข้อสอบนั้นไปทดลองใช้แล้ว โดยนำข้อมูลที่ได้มาวิเคราะห์ทางสถิติ ซึ่งได้มีผู้เสนอวิธีการตรวจสอบข้อสอบที่ลำเอียงไว้หลายวิธีด้วยกัน สามารถสรุปได้เป็น 2 แนวทาง ดังนี้

2.1 การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายนอก (External Criterion)

การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายนอกนั้น ถือได้ว่าเป็นการหาความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) ของแบบทดสอบนั่นเอง สามารถทำได้โดยการเขียนกราฟแสดงความสัมพันธ์ระหว่างคะแนนที่ใช้เป็นเกณฑ์ภายนอกกับตัวแปรทำนาย ซึ่งเกณฑ์ภายนอก ได้แก่ เกรดเฉลี่ย ผลการปฏิบัติงาน เป็นต้น ส่วนตัวแปรทำนาย หมายถึง ข้อสอบ โดยการพิจารณาค่าความลาดชันของกราฟ (Slope) และค่าจุดตัดแกน (Intercept) ของแต่ละกลุ่ม ซึ่งถ้ามีความแตกต่างกัน แสดงว่า ข้อสอบข้อนั้นมีควมลำเอียงต่อกลุ่มที่มีค่าตัดแกนหรือค่าความชันมากกว่า แต่การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายนอกนี้ มีจุดอ่อนมากในทางปฏิบัติที่สามารถทำได้ยากมากในการหาตัวแปรเกณฑ์ภายนอกที่มีความเที่ยงตรงเชิงพยากรณ์และมีความยุติธรรม ซึ่งจะส่งผลให้การตรวจสอบความลำเอียงขาดความสมบูรณ์ ดังนั้น จึงไม่ได้รับความนิยมเท่าที่ควร

2.2 การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายใน (Internal Criterion)

การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายในนี้ ยึดหลักที่ว่า ข้อสอบที่ผู้สอบที่มีระดับความสามารถเท่ากัน แต่อยู่ต่างกลุ่มกัน มีคะแนนผลการสอบที่แตกต่างกัน จะเป็นข้อสอบที่มีความลำเอียง ซึ่งการตรวจสอบความลำเอียงโดยวิธีนี้ได้รับความนิยมเป็นอย่างมาก (ญาณภัทร สีหะมงคล, 2540 : 17) เนื่องจากไม่ต้องกังวลเรื่องเกณฑ์สามารถศึกษาจากการตรวจสอบโครงสร้างภายในของข้อสอบ โดยพิจารณาคะแนนที่ได้จากผลการตอบข้อสอบของผู้เข้าสอบต่างกลุ่มกันว่า

วัดในคุณลักษณะที่ต้องการวัดในโครงสร้างเดียวกันหรือไม่ ถือเป็นกระบวนการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) ของแบบทดสอบนั่นเอง การตรวจสอบความลำเอียงโดยใช้เกณฑ์ภายในนี้ มีอยู่หลายวิธีด้วยกัน ซึ่งสามารถแบ่งได้เป็น 2 กลุ่มใหญ่ ๆ คือ

2.2.1 การตรวจสอบข้อสอบที่ลำเอียงโดยยึดทฤษฎีการวัดผลแบบดั้งเดิม (Classical Test Theory : CTT)

เป็นวิธีการตรวจสอบความลำเอียงที่นิยมใช้กันมาตั้งแต่เริ่มต้นของการพัฒนาเทคนิควิธีการตรวจสอบความลำเอียง โดยอธิบายถึงอิทธิพลของความคลาดเคลื่อนของการวัดที่มีต่อคะแนนจริงของผู้สอบ โดยอาศัยข้อตกลงเบื้องต้นและข้อสรุปบางประการที่ได้จากหลักการทางสถิติและคณิตศาสตร์ในการพิสูจน์และหาข้อสรุป ซึ่งมีหลายวิธีที่มีผู้นิยมทำการศึกษาไว้ เช่น วิธีวิเคราะห์ความแปรปรวน (Analysis of Variance) วิธีวิเคราะห์องค์ประกอบ (Factor Analysis) วิธีวิเคราะห์ด้วยไคสแควร์ (Chi-Square) และวิธีแปลงค่าความยาก (Transformed Item Difficulty) เป็นต้น

2.2.2 การตรวจสอบข้อสอบที่ลำเอียงโดยยึดทฤษฎีการตอบข้อคำถาม (Item Response Theory : IRT)

ทฤษฎีการตอบข้อคำถาม เป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างความสามารถที่อยู่ในตัวบุคคลกับพฤติกรรมการตอบข้อสอบของบุคคลนั้น โดยมีความเชื่อว่า พฤติกรรมการตอบข้อสอบของบุคคลจะถูกกำหนดโดยความสามารถที่มีอยู่ในตัวบุคคลนั่นเอง (Lord and novick, 1968 : 358) ซึ่งไม่สามารถจะสังเกตได้โดยตรง เขียนในรูปฟังก์ชันทางคณิตศาสตร์ได้ดังนี้

$$p = f(\theta)$$

เมื่อ	p	แทน	ผลการสอบ
	θ	แทน	ความสามารถ
	f	แทน	ฟังก์ชัน

ซึ่งวิธีการตรวจสอบข้อสอบที่ลำเอียงที่นิยมใช้กันคือ วิธีไค้ลักษณะข้อสอบ 1 พารามิเตอร์ วิธีไค้ลักษณะข้อสอบ 2 พารามิเตอร์ และ วิธีไค้ลักษณะข้อสอบ 3 พารามิเตอร์ เป็นต้น สำหรับการตรวจสอบความลำเอียงโดยยึดทฤษฎี IRT นี้ พบว่า เป็นวิธีการวิเคราะห์ที่ให้ผลที่ถูกต้องที่สุด แต่มีข้อจำกัดที่ว่า ต้องใช้การคำนวณที่ยุ่งยากซับซ้อน ใช้กลุ่มตัวอย่างขนาดใหญ่และสิ้นเปลืองค่า

ใช้จ่ายเป็นจำนวนมาก (Narayanan and Swaminathan, 1994 : 315) ทำให้เกิดความยุ่งยากในทางปฏิบัติ อย่างไรก็ตาม ปัจจุบันนักวัดผลได้พัฒนาเทคนิควิธีการใหม่ ๆ ขึ้นมาเพื่อใช้ในการตรวจสอบข้อสอบที่ลำเอียง เช่น วิธีการทำให้เป็นมาตรฐาน (Standardization Method) แบบจำลองการถดถอยโลจิสติก (Logistic Regression Model) รวมทั้งวิธีแมนเทล-แฮนส์เซล (Mantel-Haenszel) และ วิธีชิปเทสต์ (SIBTEST) ซึ่งเป็นวิธีที่ใช้สถิติแบบนอนพารามตริกในการวิเคราะห์ เป็นต้น ในการศึกษาครั้งนี้ผู้วิจัยได้ใช้วิธีการตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีชิปเทสต์ และวิธีแมนเทล-แฮนส์เซล ซึ่งทั้งสองวิธีนี้เป็นวิธีที่มีผู้นิยมนำมาใช้ในทางปฏิบัติ เนื่องจากมีประสิทธิภาพในการตรวจสอบ มีวิธีการคำนวณที่ไม่ยุ่งยากซับซ้อน และสิ้นเปลืองค่าใช้จ่ายไม่มาก ซึ่งสามารถวิเคราะห์ได้ด้วยคอมพิวเตอร์ส่วนบุคคล ดังมีรายละเอียดต่อไปนี้

การตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีชิปเทสต์ (SIBTEST : SIB)

การตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีชิปเทสต์นี้ เชนอ โดย ชิลลี และสตาท์ (Shealy and Stout, 1993 : 159 - 194) ซึ่งเป็นวิธีการที่มีแนวความคิดบนการตรวจสอบความลำเอียงของแบบทดสอบชนิดพหุมิติ (Multidimensional) โดยมีพื้นฐานบนทฤษฎีการตอบข้อคำถามแบบพหุมิติ แต่การทดสอบ SIB แตกต่างจาก IRT ตรงที่เป็นการใช้การทดสอบค่าสถิตินอนพารามตริก และมีการคำนวณที่ง่ายไม่ซับซ้อน ใช้ได้ดีสำหรับการตรวจสอบความลำเอียงอย่างสม่ำเสมอและมีทิศทางเดียว มีข้อตกลงว่ามีมิติการวัด 2 มิติ ๆ หนึ่งเป็นความสามารถเป้าหมายที่ต้องการวัด ใช้สัญลักษณ์ θ อีกมิติหนึ่งเป็นความสามารถแทรกซ้อนที่ไม่ต้องการวัด แทนด้วย η สำหรับนิยามของข้อสอบที่มีความลำเอียงในการทดสอบ SIB คือ ความลำเอียงของข้อสอบจะเกิดขึ้นต่อกลุ่มเปรียบเทียบที่ระดับความสามารถ θ ถ้าฟังก์ชันการตอบข้อสอบสำหรับความสามารถเป้าหมายของกลุ่มอ้างอิงมีค่ามากกว่าฟังก์ชันการตอบข้อสอบของกลุ่มเปรียบเทียบ (Shealy and Stout, 1993 : 159 - 194) ซึ่งแสดงในรูปของ

$$T_R(\theta) > T_F(\theta)$$

เมื่อ $T_R(\theta) = E_R[P(\theta, \eta)/\theta]$ โดยที่ R คือ กลุ่มอ้างอิง และ F คือ กลุ่มเปรียบเทียบ

θ คือ ความสามารถเป้าหมาย

η คือ ความสามารถสอดแทรก

$T_R(\theta)$ คือ ฟังก์ชันการตอบข้อสอบของกลุ่มอ้างอิง

$T_F(\theta)$ คือ ฟังก์ชันการตอบข้อสอบของกลุ่มเปรียบเทียบ

โดยข้อสอบจะเกิดความลำเอียง เมื่อ $\eta_F/\theta = \eta_R/\theta$ สำหรับนิยามของข้อสอบที่มีความลำเอียงในแง่ของแบบทดสอบทั้งฉบับ เมื่อให้ U เป็นการตอบข้อสอบแบบสุ่มของผู้สอบคนหนึ่ง $h(U)$ คือ คะแนนจากแบบทดสอบ โดยที่ $h(U) = \sum U_i$ จะได้ว่า ความลำเอียงของข้อสอบจะเกิดต่อกลุ่มเปรียบเทียบในระดับความสามารถ θ ถ้า

$$T_R(\theta) > T_F(\theta)$$

$$T_R(\theta) \text{ def } E_R[h(U)/\theta]$$

$$T_F(\theta) \text{ def } E_F[h(U)/\theta]$$

โดยมีสมมติฐานในการทดสอบข้อสอบที่มีความลำเอียงดังนี้

$$H_0 : \beta_U = 0$$

$$H_1 : \beta_U > 0$$

เมื่อ β_U เป็นพารามิเตอร์ที่แสดงถึงข้อสอบที่ลำเอียง โดยที่ $\beta_U = \int \beta(\theta) f_F(\theta) d(\theta)$ และ $\beta(\theta) = T_R(\theta) - T_F(\theta)$ เป็นความแตกต่างในการศึกษาฟังก์ชันการตอบข้อสอบสำหรับกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ

เมื่อให้ N แทนผลรวมของจำนวนข้อสอบ และให้ข้อสอบข้อที่ $1, \dots, n$ แทนข้อสอบที่มีความเที่ยงตรง และข้อสอบข้อที่ $n+1, \dots, N$ แทนข้อสอบที่ต้องการศึกษา ให้ U_i แทนผลการตอบข้อสอบข้อที่ i โดยมีค่าเป็น 0 ถ้าตอบผิด และมีค่าเป็น 1 ถ้าตอบถูก โดยที่คะแนนของแต่ละแบบทดสอบ คือ

$$X = \sum_{i=1}^n U_i \quad \text{แทน คะแนนรวมของแบบทดสอบที่มีความเที่ยงตรง}$$

$$\text{และ } Y = \sum_{i=n+1}^N U_i \quad \text{แทน คะแนนรวมของแบบทดสอบที่ต้องการศึกษา}$$

สำหรับค่าสถิติที่ประมาณค่าความลำเอียงคำนวณได้จาก

$$\hat{\beta}_U = \sum_{k=0}^n \hat{p}_k (\bar{Y}_{Rk} - \bar{Y}_{Fk})$$

เมื่อ $\hat{p}_k$ แทนสัดส่วนของผู้สอบกลุ่มเปรียบเทียบที่ตอบแบบทดสอบย่อยที่ถือว่าไม่
 ลำเอียงถูก k ข้อ $\bar{y}_{Rk}$ และ $\bar{y}_{Fk}$ เป็นคะแนนเฉลี่ยที่ได้จากแบบทดสอบย่อยที่ต้องการศึกษา
 สำหรับผู้สอบทุกคนของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบตามลำดับที่มีคะแนนรวมในแบบทดสอบ
 ย่อยที่ถือว่าเป็นแบบทดสอบที่มีความเที่ยงตรง โดยกำหนดสถิติทดสอบ SIBTEST คือ

$$B_U = \hat{\beta}_U / \sigma(\hat{\beta}_U)$$

เมื่อ $\sigma(\hat{\beta}_U)$ เป็นการประมาณค่าความคลาดเคลื่อนมาตรฐานของ $(\hat{\beta}_U)$ และ

$$\sigma(\hat{\beta}_U) = \left\{ \sum_{k=0}^n \hat{p}_k^2 \left[\frac{1}{j_{Rk}} \sigma^2(y/R, k) + \frac{1}{j_{Fk}} \sigma^2(y/F, k) \right] \right\}^2$$

โดยที่ j_{Rk}, j_{Fk} เป็น จำนวนผู้สอบในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่ทำคะแนน
 ได้เท่ากัน k ในแบบทดสอบย่อยที่เที่ยงตรง ($X = k$)

การตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีของแมนเทล-แฮนส์เซล (Mantel-Haenszel : MH)

การตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีของแมนเทล-แฮนส์เซลนี้ แมนเทลและแฮนส์เซล
 ได้เสนอขึ้นใช้ตั้งแต่ปี 1959 ซึ่งฮอลแลนด์และเทเยอร์ (Holland and Thayer) ได้นำมาใช้ในการ
 ตรวจสอบความลำเอียงในปี 1988 ซึ่งเทคนิค MH เป็นการเปรียบเทียบผลการตอบข้อสอบของกลุ่ม
 ผู้สอบ 2 กลุ่ม คือกลุ่มอ้างอิง (Reference Group) และกลุ่มเปรียบเทียบ (Focal Group) โดยกลุ่มแรก
 จะใช้อ้างถึงกลุ่มที่คาดว่า จะได้ประโยชน์จากข้อสอบและกลุ่มหลังเป็นกลุ่มที่เสียประโยชน์จาก
 ข้อสอบ ในกรณีที่ข้อสอบมีความลำเอียงโดยมีการตรวจสอบทุก ๆ ระดับคะแนนรวมจากการสอบ
 ข้อสอบใดที่ผู้สอบทั้งสองกลุ่มทำได้เท่า ๆ กัน จะเป็นข้อสอบที่ถือว่า ไม่มีความลำเอียงต่อกลุ่ม
 ใดกลุ่มหนึ่ง วิธีการในการตรวจสอบมีดังนี้

ถ้าให้ N_{Rj} แทนจำนวนผู้สอบในกลุ่มอ้างอิง และ N_{Fj} แทนจำนวนผู้สอบในกลุ่มเปรียบเทียบ
 ที่มีคะแนนรวมในช่วงคะแนน j และ ให้ $N_j = N_{Rj} + N_{Fj}$ เป็นจำนวนผู้สอบรวมทั้ง 2 กลุ่มที่ได้
 คะแนนรวมจากการสอบข้อสอบข้อที่ i ในช่วงคะแนน j นำมาเขียนในรูปตาราง 2 ทาง แสดงผล
 การตอบถูก (1) และผิด (0) ดังแสดงในตาราง 1 (Narayanan and Swaminathan. 1994 : 315 - 317)

ตาราง 1 ข้อมูลการสอบของกลุ่มตัวอย่างกลุ่ม R และ กลุ่ม F ที่มีคะแนนรวมอยู่ในช่วงคะแนน j ของข้อสอบข้อที่ i

กลุ่ม	คะแนนที่ได้จากข้อสอบที่ต้องการตรวจสอบ		
	1	0	รวม
กลุ่มอ้างอิง (R)	A_j	B_j	N_{Rj}
กลุ่มเปรียบเทียบ (F)	C_j	D_j	N_{Fj}
รวม	N_{1j}	N_{0j}	N_j

จากตาราง 1 จะได้ว่า ในช่วงระดับคะแนน j A_j และ B_j คือ จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบข้อที่ถูกต้องและไม่ถูกต้อง ตามลำดับ ในขณะที่ C_j และ D_j คือจำนวนผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบข้อที่ถูกต้องและไม่ถูกต้อง ตามลำดับ โดยที่ A_j , B_j , C_j และ D_j เป็นความถี่ที่สังเกตได้ของการตอบถูก (1) และผิด (0) ในช่วงคะแนนที่ j ซึ่งข้อมูลการตอบข้อสอบแต่ละข้อที่ถูกนำมาแสดงในรูปตาราง 2 ทางนี้ เป็นจำนวนตารางเท่ากับคะแนนรวมที่มีผู้สอบทำได้นับด้วย 1 จากนั้นจะคำนวณค่าความน่าจะเป็นในรูปของค่าสัดส่วนการตอบข้อสอบถูก-ผิดระหว่างกลุ่ม โดยกำหนดให้เป็นค่า α_{MH} ดังนี้

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

ค่า α_{MH} จะมีค่าระหว่าง 0 และ ∞ โดยที่ $\alpha_{MH} = 1$ จะแสดงถึงสมมติฐานที่เป็นกลาง คือ ไม่มีความลำเอียง ซึ่งแสดงว่า กลุ่มตัวอย่างทั้งสองกลุ่มทำข้อสอบได้ผลเช่นเดียวกันในแต่ละระดับคะแนนรวม ค่า $A_j D_j / N_j$ และ $B_j C_j / N_j$ มีค่าเท่ากันทั้งสองกลุ่ม ทำให้ $\alpha_{MH} = 1$ แต่ถ้า α_{MH} มีค่ามากกว่า 1 แสดงว่า กลุ่มอ้างอิงจะทำข้อสอบข้อนั้นได้ถูกมากกว่ากลุ่มเปรียบเทียบ และถ้า α_{MH} มีค่าน้อยกว่า 1 แสดงว่ากลุ่มเปรียบเทียบจะทำข้อสอบข้อนั้นได้ถูกมากกว่ากลุ่มอ้างอิง สำหรับสมมติฐานที่เป็นกลางที่แสดงว่าข้อสอบไม่ลำเอียง เป็นดังนี้

$$H_0 : [\pi_{Rj} / (1 - \pi_{Rj})] = [\pi_{Fj} / (1 - \pi_{Fj})]$$

$$H_1 : [\pi_{Rj} / (1 - \pi_{Rj})] = \alpha [\pi_{Fj} / (1 - \pi_{Fj})]$$

เมื่อ $j = 1, 2, 3, \dots, k, \alpha \neq 1$

และ π_{Rj} เป็น ค่าความน่าจะเป็นที่กลุ่มอ้างอิงตอบข้อสอบข้อนั้นถูกในระดับ
ช่วงคะแนน j

π_{Fj} เป็น ค่าความน่าจะเป็นที่กลุ่มเปรียบเทียบตอบข้อสอบข้อนั้นถูกใน
ระดับช่วงคะแนน j

อย่างไรก็ตาม ในทางปฏิบัติแล้ว แม้จะมีการควบคุมให้ผู้สอบ 2 กลุ่มมีความสามารถ
ในสิ่งที่ต้องการวัดเท่ากัน แต่เป็นไปได้ที่ผลการตอบข้อสอบของผู้สอบทั้งสองกลุ่ม จะเป็นเช่น
เดียวกันดังที่กล่าว แมนเทิลและแฮนสเซล จึงได้เสนอค่าสถิติไคสแควร์เพื่อทดสอบค่าที่ได้ว่าจะมี
ความแตกต่างจาก 1.00 อย่างมีนัยสำคัญหรือไม่ ที่ $df = 1$ ดังนี้

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

โดยที่ค่าผลรวมเป็นค่าที่ได้จากการรวมทุกชั้นของคะแนนรวม ซึ่ง $E(A_j), Var(A_j)$ มี
ค่าดังนี้

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

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

เมื่อ $E(A_j)$ และ $Var(A_j)$ เป็นค่าคาดหวังและความแปรปรวนของ A ที่ระดับคะแนน j

นอกจากนี้ Education Testing Service ได้เสนอให้แปลงค่า α_{MH} ให้เป็นคะแนนมาตรฐาน
ในรูปเคลด้าที่มีค่าเฉลี่ยเป็น 0 โดยเรียกค่าที่แปลงนี้ว่า MH D. DIF ที่มีสมการดังนี้ (Millsap and
Everson, 1993 : 303)

$$\alpha_{MH} = \frac{-4}{1.7} \ln(\alpha_{MH}) = -2.35 \ln(\alpha_{MH})$$

ค่า Δ_{MH} มีค่าระหว่าง -2.6 ถึง 2.6 และใช้เป็นดัชนีแสดงความลำเอียง ข้อสอบที่ไม่มี
ความลำเอียงจะมีค่า $\alpha_{MH} = 1$ หรือ $\Delta_{MH} = 0$

3. หลักการวิเคราะห์ค่าความง่ายของข้อสอบ

การวิเคราะห์ค่าความง่ายของข้อสอบ เป็นการหาคุณภาพรายข้อของแบบทดสอบ โดยนำผลที่ได้จากการสอบของนักเรียนมาทำการวิเคราะห์หาคุณภาพเป็นรายข้อ โดยการพิจารณาจากดัชนีค่าความง่ายของข้อสอบ (Easiness Index) ซึ่งเป็นค่าที่แสดงคุณสมบัติของข้อสอบว่า ในข้อสอบข้อหนึ่ง ๆ มีนักเรียนตอบถูกกี่คนจากจำนวนนักเรียนทั้งหมดที่ทำข้อสอบข้อนั้น ถ้ามีจำนวนนักเรียนทำถูกมากก็จะมีค่าดัชนีความง่ายสูง ซึ่งแปลว่าเป็นข้อสอบที่ง่ายมาก แต่ถ้ามีนักเรียนทำถูกน้อย ก็จะมีค่าดัชนีความง่ายต่ำ ซึ่งแปลว่า เป็นข้อสอบที่ง่ายน้อยหรือเป็นข้อสอบที่ยากมากนั่นเอง (ล้วน สายยศ และอังคณา สายยศ. 2539 :184 - 185) โดยคำนวณจากสัดส่วนของนักเรียนที่ทำข้อสอบข้อนั้นถูก ซึ่งมีสูตรในการคำนวณ ดังนี้

$$p_E = \frac{N_r}{N_t}$$

เมื่อ	p_E	แทน	ดัชนีค่าความง่าย
	N_r	แทน	จำนวนนักเรียนที่ทำข้อสอบข้อนั้นถูก
	N_t	แทน	จำนวนนักเรียนทั้งหมดที่ทำข้อสอบข้อนั้น

การคำนวณค่าความง่ายโดยทั่วไป ข้อสอบที่เหมาะสมจะมีดัชนีค่าความง่าย (p_E) เท่ากับ 0.50 และในการคัดเลือกข้อสอบนั้น จะพิจารณาค่า p_E อยู่ระหว่าง 0.20 – 0.80 จึงจะถือว่าเป็นข้อสอบที่ใช้ได้ กล่าวคือ ยิ่งมีค่า p_E สูง จะแปลว่าข้อสอบข้อนั้นง่าย และค่า p_E ต่ำ แปลว่าข้อสอบนั้นยากนั่นเอง ซึ่งการแปลความหมายค่า p_E จะแปลผลดังตารางต่อไปนี้

ตาราง 2 การแปลความหมายของดัชนีค่าความง่าย (p_E) ของข้อสอบ

ดัชนีค่า p_E	ความหมาย
มากกว่า 0.80	ง่ายมาก (ปรับปรุงหรือตัดทิ้ง)
0.60 – 0.80	ค่อนข้างง่าย
0.40 – 0.59	ปานกลาง
0.20 – 0.39	ค่อนข้างยาก
ต่ำกว่า 0.20	ยากมาก(ปรับปรุงหรือตัดทิ้ง)

4. งานวิจัยที่เกี่ยวข้อง

ก. งานวิจัยในประเทศ

จากการศึกษางานวิจัยภายในประเทศที่ผ่านมา ได้มีผู้ทำการวิจัยเกี่ยวกับการตรวจสอบข้อสอบที่ลำเอียงไว้หลายท่านด้วยกัน โดยศึกษาในรูปแบบการเปรียบเทียบผลการตรวจสอบความลำเอียงด้วยวิธีต่าง ๆ กับกลุ่มย่อยที่มีลักษณะแตกต่างกันในด้านภาษา เพศ และ สภาพภูมิศาสตร์ รวมทั้งการเปรียบเทียบคุณภาพของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว นอกจากนี้ ยังมีการพัฒนาเกณฑ์ที่ใช้ในการตัดสินความลำเอียงทางเพศ ซึ่งได้นำเสนอดังต่อไปนี้

ซัชชัย เผ่าพงษ์ (2527 : 73 - 76) ได้ทำการตรวจสอบความลำเอียงของข้อสอบมาตรฐานวัดความถนัดทางการเรียน ด้านคณิตศาสตร์และภาษา ในระดับมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ระหว่างกลุ่มผู้สอบเพศชายและเพศหญิง โดยการเปรียบเทียบโค้งลักษณะข้อสอบด้วยวิธีแบบจำลองโลจิสติกแบบ 3 พารามิเตอร์ กลุ่มตัวอย่างที่ใช้ในการศึกษา เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2534 เพศชายและเพศหญิงซึ่งประกอบด้วยทุกภาคภูมิศาสตร์ทั่วประเทศ 5 ภาค คือ ภาคเหนือ ภาคกลาง ภาคใต้ ภาคตะวันออก และภาคตะวันออกเฉียงเหนือ กลุ่มซึ่งทำการวิเคราะห์ด้วยแบบทดสอบวัดความถนัดทางการเรียนด้านคณิตศาสตร์เป็นนักเรียนชาย 1,610 คน และนักเรียนหญิง 1,337 คน ส่วนอีกกลุ่มหนึ่ง เป็นนักเรียนชาย 1,316 คน และนักเรียนหญิง 985 คน ทำการวิเคราะห์ด้วยแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ผลการวิเคราะห์ พบว่า จากแบบทดสอบวัดความถนัดทางการเรียนด้านคณิตศาสตร์ จำนวน 30 ข้อ มีข้อสอบที่ลำเอียงต่อกลุ่มใดกลุ่มหนึ่ง เฉพาะ 9 ข้อ คือ ลำเอียงต่อกลุ่มนักเรียนชาย 7 ข้อ และลำเอียงต่อกลุ่มนักเรียนหญิง 2 ข้อ ข้อสอบที่ลำเอียงต่อกลุ่มนักเรียนชายและนักเรียนหญิงในระดับปานกลางขึ้นไป มีจำนวน 5 ข้อ ซึ่งวัดในเรื่อง ร้อยละ การหาปริมาตร และ การหาความยาวของเส้นในด้านของรูปสามเหลี่ยม จำนวนเรื่องละ 1 ข้อ และเรื่องเกี่ยวกับโจทย์ปัญหา จำนวน 2 ข้อ จากข้อสอบจำนวน 5 ข้อดังกล่าว เป็นข้อสอบที่ลำเอียงต่อกลุ่มนักเรียนหญิงในช่วงความสามารถแรก และลำเอียงต่อกลุ่มนักเรียนชายในช่วงความสามารถต่อมา จำนวน 1 ข้อ ส่วนแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จำนวน 40 ข้อ มีข้อสอบที่ลำเอียงต่อนักเรียนชายโดยเฉพาะ 3 ข้อ และลำเอียงต่อนักเรียนหญิงโดยเฉพาะ 8 ข้อ ข้อสอบที่ลำเอียงในระดับปานกลางขึ้นไป มีจำนวน 9 ข้อ ซึ่งวัดความเข้าใจเกี่ยวกับการอ่านคำ ประพันธ์ บทร้อยกรองอย่างละ 1 ข้อ และวัดความเข้าใจเกี่ยวกับการอ่านข้อความ จำนวน 7 ข้อ จากข้อสอบจำนวน 9 ข้อขึ้นไป เป็นข้อสอบที่ลำเอียงต่อนักเรียนชายโดยเฉพาะ 1 ข้อ และเป็น

ข้อสอบที่ลำเอียงต่อนักเรียนหญิงโดยเฉพาะจำนวน 6 ข้อ และมีข้อสอบจำนวน 2 ข้อ มีความลำเอียงทางการทดสอบมากกว่า 1 ช่วงความสามารถ

ทัศนีย์ พิรมนตรี (2530 : 7) ได้ทำการตรวจสอบความลำเอียงของแบบทดสอบวิชาคณิตศาสตร์ โครงการตรวจสอบคุณภาพการศึกษาของสำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2536 จำนวน 7,036 คน โดยเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงระหว่างนักเรียนในกรุงเทพมหานครกับนักเรียนในภาคภูมิศาสตร์ทั้ง 5 ภาค คือ ภาคกลาง ภาคตะวันออก ภาคใต้ และ ภาคตะวันออกเฉียงเหนือ ด้วยวิธีตรวจสอบความลำเอียง 3 วิธี คือ วิธีกำหนดจุดเดลด้า วิธีทดสอบความแตกต่างระหว่างกลุ่มด้วยสถิติไคสแควร์ในโมเดลล็อกลิเนีย 2 โมเดล คือ โมเดลที่ไม่มีค่าพารามิเตอร์ผลรวมระหว่างระดับคะแนนกับกลุ่ม และโมเดลที่ไม่มีพารามิเตอร์ของผลหลักที่เกิดจากกลุ่ม และวิธีโค้งลักษณะข้อสอบที่ใช้พารามิเตอร์ 3 ตัว พบข้อสอบที่มีความลำเอียงจากวิธีโค้งลักษณะข้อสอบที่ใช้พารามิเตอร์ 3 ตัวมากที่สุด และในแต่ละวิธีตรวจสอบความลำเอียง พบว่า มีข้อสอบที่ลำเอียงซ้ำกันระหว่างกลุ่มนักเรียนในกรุงเทพมหานคร กับนักเรียนในภาคภูมิศาสตร์ทุกภาค

นิรมล ชัยขวลิต (2537 : 57 - 58) ได้ศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงของแบบทดสอบวัดความเข้าใจในการอ่านภาษาไทย จำนวน 50 ข้อ ระหว่างกลุ่มผู้สอบที่มีเพศต่างกัน ด้วยวิธีตรวจสอบความลำเอียงตามทฤษฎีคลาสสิกอล 3 วิธี คือ วิธีแปลงค่าความยาก วิธีไคสแควร์ และวิธีวิเคราะห์ความแปรปรวน โดยศึกษาเปรียบเทียบจำนวนข้อสอบที่ลำเอียง ความเชื่อมั่นของแบบทดสอบหลังคัดข้อที่ลำเอียงออกแล้ว ที่คำนวณด้วยสูตร KR-20 และ แบบแบ่งครึ่งฉบับ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษาส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวน 1,006 คน ซึ่งได้มาโดยการสุ่มอย่างง่าย ผลการศึกษา พบว่า การตรวจสอบความลำเอียงด้วยวิธีแปลงค่าความยากกับวิธีไคสแควร์ และวิธีแปลงค่าความยากกับวิธีวิเคราะห์ความแปรปรวน มีจำนวนข้อสอบที่ลำเอียงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีไคสแควร์กับวิธีวิเคราะห์ความแปรปรวน มีจำนวนข้อสอบที่ลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าความเชื่อมั่นของแบบทดสอบหลังคัดข้อสอบที่ลำเอียงออกแล้วที่คำนวณด้วยสูตร KR-20 และสูตรแบบแบ่งครึ่งฉบับที่วิเคราะห์ด้วยวิธีทั้ง 3 พบว่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

จิตสุดา ธารพร (2539 : 128 - 130) ได้ศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงของข้อสอบจากแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ และ แบบทดสอบความเข้าใจทางภาษา ของนักเรียนที่นับถือศาสนาพุทธและศาสนาอิสลาม โดยใช้วิธีตรวจสอบความลำเอียงต่างกัน 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบ และ วิธีของแมนเทล-แฮนส์เซล เพื่อ

เปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง และเปรียบเทียบค่าความเชื่อมั่นที่คำนวณโดยใช้สูตรแบบ KR-20 และ ใช้สูตรแบบแบ่งครึ่งฉบับของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ของจังหวัดกระบี่ จำนวน 2,689 คน ซึ่งเลือกมาโดยการสุ่มอย่างง่าย

ผลการวิจัย พบว่า วิธีตรวจสอบความลำเอียงของข้อสอบจากแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์และแบบทดสอบความเข้าใจทางภาษาแต่ละวิธี พบจำนวนข้อสอบที่มีความลำเอียงมากที่สุด และวิธีแปลงค่าความยากของข้อสอบพบจำนวนข้อสอบที่มีความลำเอียงน้อยที่สุด โดยที่การตรวจสอบความลำเอียงด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธี โคง์ลักษณะข้อสอบ และวิธีแปลงค่าความยากของข้อสอบกับวิธีของแมนเทิล-แฮนส์เชล มีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธี โคง์ลักษณะข้อสอบกับวิธีของแมนเทิล-แฮนส์เชล มีจำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับค่าความเชื่อมั่นของแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์ที่คำนวณโดยใช้สูตรแบบ KR-20 หลังการคัดเลือกข้อสอบที่มีความลำเอียงออกจากการตรวจสอบความลำเอียงด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธี โคง์ลักษณะข้อสอบ และวิธีแปลงค่าความยากกับวิธีของแมนเทิล-แฮนส์เชล มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และวิธี โคง์ลักษณะข้อสอบกับวิธีของแมนเทิล-แฮนส์เชล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นของแบบทดสอบที่คำนวณโดยใช้สูตรแบบแบ่งครึ่งฉบับ หลังการคัดเลือกข้อสอบที่มีความลำเอียงออกจากการตรวจสอบความลำเอียงด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธี โคง์ลักษณะข้อสอบ และวิธีของแมนเทิล-แฮนส์เชล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบความเข้าใจทางภาษาที่คำนวณโดยใช้สูตร KR-20 และสูตรแบบแบ่งครึ่งฉบับ หลังการคัดเลือกข้อสอบที่มีความลำเอียงออกจากการตรวจสอบความลำเอียงด้วยวิธีแปลงค่าความยากของข้อสอบกับวิธี โคง์ลักษณะข้อสอบ และวิธีของแมนเทิล-แฮนส์เชล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

คมศักดิ์ ชื่นชม (2539 : 108) ได้ทำการศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงของแบบทดสอบวัดจริยธรรมความซื่อสัตย์ ตามทฤษฎีคลาสสิกอลที่ใช้วิธีต่างกัน 3 วิธี คือ วิธีค่าอำนาจจำแนก วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอย โดยศึกษาเปรียบเทียบจำนวนข้อสอบที่ลำเอียง และค่าความเชื่อมั่นของแบบทดสอบหลังคัดข้อที่ลำเอียงออกแล้ว กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2538 ในจังหวัดยะลา จำนวน 1,081 คน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น ผลการศึกษา พบว่า การตรวจสอบความลำเอียงของด้วยวิธีค่าอำนาจจำแนก วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอย มีจำนวนข้อสอบที่ลำเอียงน้อย

และไม่แตกต่างกัน ที่ระดับนัยสำคัญ .05 ส่วนค่าความเชื่อมั่นของแบบทดสอบวัดจริยธรรมความซื่อสัตย์ หลังคัดข้อที่ลำเอียงออกแล้ว จากการตรวจสอบความลำเอียงด้วยวิธีค่าอำนาจจำแนก วิธีวิเคราะห์องค์ประกอบ และวิธีวิเคราะห์การถดถอย มีค่าไม่แตกต่างกัน ที่ระดับนัยสำคัญ .05

เรวดี อินทสระ (2539 : 248 - 252) ได้ทำการศึกษาความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกรับรองตรวจสอบความลำเอียงต่อเพศตามทฤษฎีการตอบข้อคำถาม (IRT) วิธีแมนเทิล-แฮนส์เซล (MH) และวิธีซิปเทสท์ (SIBTEST) การตัดสินผลการสอบที่คิดคะแนนมาตรฐานที่-ปกติ และคะแนนน้ำหนักรับรองความสามารถและสาเหตุความลำเอียงของข้อสอบ โดยศึกษาความลำเอียงของข้อสอบคัดเลือกรับรองเข้าศึกษาในชั้นปีที่ 1 ประเภทรับตรง ปีการศึกษา 2538 ของมหาวิทยาลัยสงขลานครินทร์ ในวิชาภาษาไทย ก วิชาสังคมศึกษา ก วิชาภาษาอังกฤษ กข วิชาละ 8,127 คน วิชาภาษาไทย กข วิชาสังคมศึกษา กข และวิชาภาษาอังกฤษ กขค วิชาละ 5,415 คน ความเที่ยงตรงเชิงพยากรณ์ศึกษา จากคะแนนสอบคัดเลือกรับรองระดับผลการศึกษาระดับที่ 1 ปีการศึกษา 2538 ของนักศึกษาที่ได้รับการคัดเลือกจากประเภทรับตรง สายวิทยาศาสตร์ 763 คน และสายศิลปศาสตร์ 281 คน และสาเหตุความลำเอียงของข้อสอบจากการระบุสาเหตุของนักวัดผลการศึกษาหรืออาจารย์ผู้สอนจำนวน 50 คน และ นิสิตนักศึกษาที่เรียนในสาขาวิชานั้นๆ วิชาละ 50 คน

ผลการศึกษาพบว่า วิธีตรวจสอบความลำเอียงทั้ง 3 วิธี ตัดสินจำนวนข้อสอบที่ลำเอียงต่างกัน ในวิชาภาษาไทย ก ฉบับที่ 2 และวิชาสังคมศึกษา ก ฉบับที่ 1 ที่ระดับนัยสำคัญทางสถิติ .05 นอกนั้นต่างกันที่ระดับนัยสำคัญทางสถิติ .01 โดยทฤษฎีการตอบข้อคำถามตัดสินจำนวนข้อสอบที่ลำเอียงได้มากที่สุด ความสัมพันธ์ระหว่างลำดับที่ของการสอบไม่ว่าจะคิดคะแนนมาตรฐานที่-ปกติ หรือคะแนนน้ำหนักรับรองความสามารถและใช้ข้อสอบจำนวนทั้งหมดหรือใช้ข้อสอบที่ปราศจากความลำเอียงต่างมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 ประสิทธิภาพในการทำนายผลสัมฤทธิ์ทางการเรียนในสายวิทยาศาสตร์ การคิดคะแนนน้ำหนักรับรองความสามารถและใช้ข้อสอบจำนวนทั้งหมด หรือใช้ข้อสอบที่ปราศจากความลำเอียง มีประสิทธิภาพในการทำนายสูงกว่าการคิดคะแนนมาตรฐานที่-ปกติ ส่วนสายศิลปศาสตร์ การคิดคะแนนมาตรฐานที่-ปกติ ทั้งใช้ข้อสอบจำนวนทั้งหมด หรือใช้ข้อสอบที่ปราศจากความลำเอียง มีประสิทธิภาพในการทำนายสูงกว่าการคิดคะแนนความสามารถ และสาเหตุของความลำเอียงของข้อสอบต่อทั้งเพศชาย และเพศหญิง เกิดจากข้อสอบที่เป็นคำถามที่ผู้สอบได้รับการฝึกฝนเฉพาะ จะมีโอกาสตอบถูกมากกว่าเป็นเรื่องราวที่กลุ่มนั้นๆ สนใจและเป็นข้อสอบถามความจำ

อังคณา สายยศ (2539 : 24 - 26) ได้ศึกษาผลการตรวจสอบความลำเอียงของแบบทดสอบวัดความเข้าใจในการอ่านตามลักษณะการใช้ภาษาพูดในครอบครัวเป็นภาษาลาว เขมร และ ภาษาส่วย โดยศึกษาเปรียบเทียบจำนวนข้อสอบที่ลำเอียง ความเชื่อมั่นของแบบทดสอบหลังคัดข้อที่ลำเอียงออกแล้ว และคะแนนที่ได้จากการทดสอบหลังคัดข้อที่ลำเอียงออกแล้ว กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดสุรินทร์ ได้มาโดยการสุ่มแบบแบ่งชั้น มีจำนวนนักเรียนทั้งสิ้น 860 คน ทำการตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล ผลการวิจัยพบว่า จำนวนข้อสอบที่ลำเอียงของแบบทดสอบวัดความเข้าใจในการอ่านระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดในครอบครัวเป็นภาษาลาว เขมร และภาษาส่วยแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับความเชื่อมั่นของแบบทดสอบ หลังจากคัดข้อสอบที่ลำเอียงออกแล้ว พบว่า ความเชื่อมั่นระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดในครอบครัวเป็นภาษาลาว เขมร และภาษาส่วยแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนคะแนนที่ได้จากการทดสอบหลังคัดข้อที่ลำเอียงออกแล้ว มีคะแนนระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาวกับภาษาส่วย และภาษาเขมรกับภาษาส่วย แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนภาษาลาวกับภาษาเขมร แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

พรรณี จินตมาศ (2540 : 90) ได้ทำการศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงของแบบทดสอบคณิตศาสตร์ โดยวิธีวิเคราะห์ต่างกัน 3 วิธี คือ วิธีของแมนเทิล-แฮนส์เซล วิธีชิปเทสต์ และวิธีแปลงค่าความยากของข้อสอบ โดยใช้ขนาดกลุ่มตัวอย่างต่าง ๆ กันผลการศึกษาพบว่า ขนาดกลุ่มตัวอย่าง 500 คน วิธีแปลงค่าความยากของข้อสอบ พบจำนวนข้อสอบที่มีความลำเอียงน้อยที่สุด วิธีแปลงค่าความยากของข้อสอบจะตรวจพบจำนวนข้อสอบที่ลำเอียงซ้ำกับวิธีของแมนเทิล-แฮนส์เซล และวิธีชิปเทสต์ทุกข้อ วิธีชิปเทสต์พบจำนวนข้อสอบที่มีความลำเอียงมากที่สุดและเมื่อวิเคราะห์จากขนาดกลุ่มตัวอย่าง 1,000 คน วิธีของแมนเทิล-แฮนส์เซลพบจำนวนข้อสอบที่มีความลำเอียงมากที่สุด วิธีแปลงค่าความยากของข้อสอบไม่พบข้อสอบที่ลำเอียง โดยจำนวนข้อสอบที่วิเคราะห์ด้วยวิธีแปลงค่าความยากกับวิธีของแมนเทิล-แฮนส์เซล และวิธีแปลงค่าความยากกับวิธีชิปเทสต์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 นอกนั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับจำนวนข้อสอบที่วิเคราะห์ด้วยวิธีแปลงค่าความยาก ระหว่างกลุ่มผู้สอบขนาด 500 คน และขนาด 1,000 คน จะมีจำนวนข้อสอบที่มีความลำเอียง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกนั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว เมื่อวิเคราะห์จากกลุ่มตัวอย่างขนาด 500 คน จากการวิเคราะห์ความลำเอียงด้วยวิธีแปลงค่าความยากและวิธีชิปเทสต์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 นอกนั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และเมื่อวิเคราะห์จากกลุ่มตัวอย่างขนาด 1,000 คน ค่าความเชื่อมั่นของแบบทดสอบหลังการคัดเลือก

ข้อสอบที่มีความลำเอียงออกแล้วจากการตรวจสอบความลำเอียง 3 วิธีนั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ สำหรับค่าความเชื่อมั่นของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออกแล้วจากการตรวจสอบความลำเอียงด้วยวิธีชิปทดสอบระหว่างกลุ่มผู้สอบขนาด 500 คน และ 1,000 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 นอกนั้น มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

นวลอนงค์ นันทวงศ์ (2543 : 102) ได้ทำการศึกษาเปรียบเทียบความลำเอียงของข้อสอบของแบบทดสอบความเข้าใจในการอ่าน ตามลักษณะการใช้ภาษาพูดในครอบครัว โดยกลุ่มอ้างอิงเป็นกลุ่มใช้ภาษาพูดเป็นภาษาไทยและกลุ่มเปรียบเทียบเป็นกลุ่มที่ใช้ภาษาพูดเป็นภาษาลาวพวนกับภาษาลาว ซึ่งทำการวิเคราะห์ความลำเอียงด้วยวิธี 3 วิธีคือ วิธีไคสแควร์ วิธีแมนเทล-แฮนส์เซล และวิธีวิเคราะห์ความแปรปรวน โดยเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงและเปรียบเทียบค่าความเชื่อมั่นที่คำนวณด้วยสูตรแบบแบ่งครั้งฉบับของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชายและนักเรียนหญิง ในระดับชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 900 คน ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน ผลการวิจัยพบว่า ในกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาวพวน จำนวนข้อสอบที่มีความลำเอียงจากการตรวจสอบความลำเอียงระหว่างวิธีไคสแควร์กับวิธีแมนเทล-แฮนส์เซล และวิธีไคสแควร์กับวิธีวิเคราะห์ความแปรปรวน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนวิธีแมนเทล-แฮนส์เซลกับวิธีวิเคราะห์ความแปรปรวน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนในกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาว พบว่า จำนวนข้อสอบที่มีความลำเอียงจากการตรวจสอบความลำเอียงระหว่างวิธีไคสแควร์กับวิธีวิเคราะห์ความแปรปรวน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนวิธีไคสแควร์กับวิธีแมนเทล-แฮนส์เซล และวิธีแมนเทล-แฮนส์เซลกับวิธีวิเคราะห์ความแปรปรวน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จำนวนข้อสอบที่มีความลำเอียง เมื่อมีวิธีวิเคราะห์ทั้งวิธีไคสแควร์ วิธีแมนเทล-แฮนส์เซล และวิธีวิเคราะห์ความแปรปรวน ระหว่างกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาวและภาษาลาวพวน พบว่า จำนวนข้อสอบที่มีความลำเอียงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบความเข้าใจในการอ่านหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว เมื่อกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาวและภาษาลาวพวน พบว่า ต่างก็มีค่าความเชื่อมั่นระหว่างวิธีตรวจสอบความลำเอียงทั้งวิธีไคสแควร์ วิธีแมนเทล-แฮนส์เซล และวิธีวิเคราะห์ความแปรปรวน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นของแบบทดสอบความเข้าใจในการอ่านหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว เมื่อมีวิธีตรวจสอบความลำเอียงวิธีเดียวกัน เมื่อกลุ่มผู้สอบที่ใช้ภาษาพูดเป็นภาษาลาวและภาษาลาวพวน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ข. งานวิจัยในต่างประเทศ

สำหรับงานวิจัยในต่างประเทศที่เกี่ยวข้องกับการตรวจสอบข้อสอบที่ลำเอียงนั้น ได้มีผู้ทำการศึกษาไว้หลายท่าน ซึ่งผู้วิจัยได้สรุปไว้ดังนี้

ผจงจิต อินทสุวรรณ (Intasuwan, 1979 : 92) ได้ศึกษาวิธีตรวจสอบความลำเอียงที่ต่างกัน 3 วิธี คือ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ วิธีราสซ และวิธีไคสแควร์ ของแบบทดสอบ IEA กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนอายุ 14 ปี ชาวอังกฤษ อเมริกัน และนิวซีแลนด์ จำนวน 8,601 คน ผลการวิจัย พบว่า สหสัมพันธ์ระหว่างวิธีทั้ง 3 มีค่าอยู่ระหว่าง .51 - .98 และวิธีราสซ กับวิธีไคสแควร์มีค่าสหสัมพันธ์สูงถึง .98

สงบ ลักษณะ (Burrill, 1982 : 168 ; citing Laksana, 1979. *Application of Analysis of Variance Approach and ITBS From 7.*) ได้ตรวจสอบความลำเอียงของข้อสอบในแบบทดสอบ ITBS ฉบับภาษาคำศัพท์และฉบับคณิตศาสตร์ ด้วยวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีวิเคราะห์ความแปรปรวน ผลการวิจัย พบว่า วิธีวิเคราะห์ความแปรปรวนเป็นวิธีที่ใช้ได้ดีในทางปฏิบัติ แต่เป็นวิธีที่ไม่ครอบคลุม ส่วนวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ เป็นวิธีที่ครอบคลุมกว่า แต่ต้องใช้กลุ่มตัวอย่างขนาดใหญ่และการคำนวณยุ่งยากซับซ้อน

สับโคเวียต และคณะ (Subkoviak, et al. 1984 : 49 - 59) ได้ศึกษาวิธีตรวจสอบความลำเอียงของข้อสอบ 3 วิธี คือ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ วิธีไคสแควร์ และ วิธีแปลงค่าความยาก ของแบบทดสอบวัดความรู้เกี่ยวกับคำศัพท์ จำนวน 50 ข้อ เป็นคำศัพท์ภาษาอังกฤษมาตรฐาน จำนวน 40 ข้อ ศัพท์แสดงชาวผิวดำ จำนวน 10 ข้อ กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนผิวขาว จำนวน 1,021 คน เป็นนักเรียนผิวดำ จำนวน 1,008 คน ผลการวิจัยพบว่า วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ เป็นวิธีที่มีประสิทธิภาพมากที่สุด รองลงมา คือ วิธีไคสแควร์ ส่วนวิธีแปลงค่าความยากเป็นวิธีที่นำไปใช้ในทางปฏิบัติได้ ในกรณีที่ไม่มีเครื่องคอมพิวเตอร์

ฟรีเดิล และ คอสติน (Freedle and Kostin, 1990 : 329 – 343) ได้ศึกษาข้อสอบที่ลำเอียงต่อสีผิวจากแบบทดสอบ GRE (Graduate Record Examination) จำนวน 11 ฉบับ 988 ข้อ และแบบทดสอบ SAT (Scholastic Aptitude Test) จำนวน 11 ฉบับ 935 ข้อ ซึ่งเป็นแบบทดสอบวัดความถนัดด้านภาษา 4 ด้าน คือ ด้านอุปมาอุปไมย ด้านคำตรงข้าม ด้านการเติมคำในประโยคให้สมบูรณ์ และด้านความเข้าใจในการอ่าน กลุ่มตัวอย่างที่ใช้ในการศึกษาแบ่งเป็น 2 กลุ่ม คือ กลุ่มเปรียบเทียบและกลุ่มอ้างอิง กลุ่มเปรียบเทียบ คือ กลุ่มคนผิวดำที่ตอบว่าใช้ภาษาอังกฤษเป็นภาษาหลักในแบบทดสอบ GRE และแบบทดสอบ SAT จำนวน 1,688 คน และ 6,332 คน ตามลำดับ ส่วนกลุ่มอ้างอิง คือ กลุ่มคนผิวขาวที่ตอบว่าใช้ภาษาอังกฤษเป็นภาษาหลักในแบบทดสอบ GRE และแบบทดสอบ

SAT จำนวน 23,375 คน และ 34,208 คน ตามลำดับ ทำการตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีแปลงค่าความยาก และ วิธี standardization Approach ผลการศึกษาพบว่า ค่าความยากของข้อสอบสามารถนำมาใช้เป็นดัชนีบ่งชี้ขนาดของความลำเอียงของข้อสอบได้ในแบบทดสอบทั้งสองฉบับ โดยกลุ่มผู้สอบผิวดำจะทำข้อสอบที่มีค่าความยากสูงได้ดีกว่ากลุ่มผู้สอบผิวขาว แต่จะทำข้อสอบที่ง่ายได้น้อยกว่าในระดับความสามารถเดียวกัน เนื่องจากในข้อสอบที่ง่ายมักใช้คำที่มีความหมายหลายอย่าง จึงมีผลต่อค่าความยากและเป็นแหล่งที่ก่อให้เกิดความลำเอียง หลังจากนั้นใช้วิธีปรับแก้ 3 แบบในการคำนวณหาขนาดของความลำเอียง กับ การคำนวณหาขนาดความลำเอียงโดยไม่ใช้วิธีปรับแก้ ซึ่งพบว่า ให้ผลไม่แตกต่างกัน

มิลเลอร์ และ โอชิม่า (Miller and Oshima. 1992 : 381 – 388) ได้ศึกษาผลของขนาดของกลุ่มตัวอย่าง จำนวนข้อสอบที่มีความลำเอียง และ ขนาดของความลำเอียงของข้อสอบที่ประมาณค่าโดยวิธีสองขั้นตอน ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลที่จำลองขึ้นแบบสองพารามิเตอร์ จำนวน 40 ข้อ มีค่าอำนาจจำแนกเฉลี่ยเป็น 1.13 ซึ่งมีค่าระหว่าง 0.41 ถึง 3.82 ส่วนเบี่ยงเบนมาตรฐาน 0.63 มีค่าความยากเฉลี่ยเป็น 0.25 โดยมีค่าระหว่าง -1.99 ถึง 2.10 ส่วนเบี่ยงเบนมาตรฐาน 1.05 โดยมีเงื่อนไขในการตรวจสอบความลำเอียง 24 เงื่อนไข คือ พารามิเตอร์ของข้อสอบแบบสองพารามิเตอร์ ขนาดของความลำเอียงมี 3 ขนาด คือ น้อย ปานกลาง และ แบบผสม โดยที่ขนาดน้อย จะมีค่าความยากเพิ่มขึ้น 0.20 แบบปานกลางจะมีค่าความยากเพิ่มขึ้น 0.35 และแบบผสมจะมีค่าความยากเพิ่มขึ้นเป็นแบบผสมระหว่างสองแบบ เงื่อนไขของจำนวนข้อสอบที่ลำเอียง 4 แบบ คือ ร้อยละ 5 , 10 , 20 และ 40 ของแบบทดสอบ โดยตรวจสอบความลำเอียงจากทฤษฎีการตอบข้อสอบ 5 วิธี คือ SA , UA , SSOS , USOS , WSSOS และ วิธีแมนเทิล-แฮนส์เชล ในขั้นตอนที่สอง เป็นการหาจำนวนข้อสอบที่ลำเอียงที่ผิดพลาดทางบวก นั่นคือ ตรวจสอบข้อสอบที่ไม่ลำเอียงว่าเป็นข้อสอบที่ลำเอียง ซึ่งพบว่า วิธีแมนเทิล-แฮนส์เชลจะตรวจสอบพบความผิดพลาดทางบวกได้น้อย และตรวจสอบความลำเอียงได้ดีพอ ๆ กับวิธีของทฤษฎีการตอบข้อสอบ

เมเซอร์ คลอสเซอร์ และแฮมเบิลตัน (Mazor , Clauser and Hambleton. 1992 : 443 - 451) ได้ศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อการตรวจสอบความลำเอียงด้วยวิธี MH โดยใช้ข้อมูลที่จำลองด้วยโปรแกรม DATAGEN แบบ 3 พารามิเตอร์ ขนาดกลุ่มตัวอย่างที่ศึกษามี 5 ขนาด คือ 2,000 คน, 1,000 คน, 500 คน, 200 คน และ 100 คน ความยาวของแบบทดสอบชุดละ 75 ข้อ ผลการวิจัยพบว่า เมื่อใช้ขนาดกลุ่มตัวอย่าง 2,000 คน วิธี MH จะตรวจค้นพบความลำเอียงได้ผิดพลาด ร้อยละ 30 และ เมื่อใช้ขนาดกลุ่มตัวอย่างน้อยกว่า 500 คน วิธี MH จะตรวจค้นพบความลำเอียงได้ผิดพลาด ร้อยละ 50 ข้อสอบที่ไม่สามารถตรวจพบความลำเอียงได้เป็นข้อสอบที่ยากมาก ข้อสอบที่มีค่าอำนาจจำแนกต่ำ และข้อสอบที่มีค่าความยากต่างกันเล็กน้อย 2 กลุ่ม

รูสโซ และสตาท์ (Roussos and Stout. 1996 : 215 - 230) ได้ศึกษาผลการใช้จำนวนกลุ่มตัวอย่างขนาดเล็ก ความแตกต่างของค่าพารามิเตอร์ข้อสอบ รวมทั้งความแตกต่างในค่าเฉลี่ยของความสามารถของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ เพื่อศึกษาผลที่มีต่อค่าความคลาดเคลื่อนแบบที่ 1 (Type I Error) คือการยอมรับว่า ข้อสอบลำเอียงทั้งที่ไม่ลำเอียงโดยวิธี MH และวิธี SIBTEST

การศึกษาที่ 1 ใช้กลุ่มตัวอย่างเป็นกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ กลุ่มละ 100, 200, 500 และ 1,000 คน เลือกข้อสอบ 1 ข้อ จากแบบทดสอบ ASVAB เลือกข้อที่มีพารามิเตอร์ของข้อสอบ คือ a , b และ c ใกล้เคียงกับค่าเฉลี่ยของแบบทดสอบมากที่สุด คือ 1.32, 0.03 และ 0.25 (ค่าเฉลี่ย 1.22, 0.09 และ 0.02) และเป็นข้อสอบที่ไม่ลำเอียงความสามารถ (θ) ของกลุ่มตัวอย่างทั้ง 2 กลุ่มถูกจำลองให้มีการกระจายเป็นโค้งปกติ มีค่าความแปรปรวนเป็น 1.00 ความแตกต่างระหว่างค่าเฉลี่ยของการแจกแจงทั้ง 2 กลุ่มถูกจำลองให้มีค่า 0.0, 0.5 และ 1.0 แล้วคำนวณค่าความสามารถเฉลี่ยของแต่ละกลุ่มให้จุดกลางของค่าเฉลี่ยอยู่บนระดับความสามารถยากเฉลี่ยของแบบทดสอบ จำลองการตอบ 400 ครั้ง และวิเคราะห์ 400 ครั้งในแต่ละสถานการณ์

ผลการใช้ MH และ SIBTEST ตรวจสอบความลำเอียงของข้อสอบที่คัดเลือกขึ้นมา พบว่า นอกจากกลุ่มตัวอย่างขนาด 100 คน และค่าความแตกต่างของค่าความสามารถเป็น 0.0 แล้ว อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 ของวิธีทั้งสอง แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่ในกลุ่มตัวอย่างขนาด 100 คน และค่าความแตกต่างของค่าความสามารถเป็น 0.0 จะมีอัตราการเกิดความคลาดเคลื่อนแบบที่ 1 ของวิธีทั้งสอง แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่วิธี SIB มีอัตราการเกิดความคลาดเคลื่อนแบบที่ 1 มากกว่าวิธี MH และในขนาดกลุ่มตัวอย่างไม่เกิน 1,000 คน จะมีแนวโน้มการเกิดความคลาดเคลื่อนสูงทั้งสองวิธี

การศึกษาที่ 2 ใช้แบบทดสอบชุดเดิมแต่กำหนดค่าความยากและค่าอำนาจจำแนกที่แตกต่างกันดังนี้ คือ ค่าอำนาจจำแนกเป็น .04, 1.0 และ 2.5 ค่าความยากเป็น -1.5, -0.5, 0.0, 0.5 และ 1.5 ค่าการเดาเป็น 0.20 ทำการจำลองการแจกแจงความสามารถเป็น 2 แบบ คือ 0.0 และ 1.0 นอกนั้น จำลองทุกอย่างเช่นเดียวกับการศึกษาที่ 1 แต่ใช้จำนวนผู้สอบขนาดใหญ่ คือ 500, 1,000 และ 3,000 คน ในแต่ละสถานการณ์คำนวณซ้ำ 100 ครั้ง และตรวจสอบความลำเอียงโดยใช้วิธี MH และวิธี SIB ผลการวิจัยพบว่า เมื่อค่าอำนาจจำแนกมีค่าสูง และค่าความยากต่ำ อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 จะสูงทั้งสองวิธีที่ระดับนัยสำคัญ 0.5 และวิธี MH มีอัตราการเกิดความคลาดเคลื่อนแบบที่ 1 มากกว่าวิธี SIB และขนาดกลุ่มตัวอย่างที่เพิ่มขึ้น 3,000 คน มีค่าความยากเป็น -1.5 ค่าอำนาจจำแนกเป็น 2.5 อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 มากกว่าวิธี SIB เมื่อการแจกแจงความสามารถเป็น 1.0 ค่าการเดาเป็น 0.20 วิธี SIB มีอัตราการเกิดความคลาดเคลื่อน

แบบที่ 1 ต่ำกว่าวิธี MH แต่เมื่อใช้การแจกแจงความสามารถที่ 0.0 ไม่มีการเดาแล้ว อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 ของทั้งสองวิธีไม่แตกต่างกัน

สรุปได้ว่า เมื่อใช้กลุ่มตัวอย่างขนาดเล็กไม่เกิน 1,000 คน อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 ของทั้งสองวิธีไม่แตกต่างกัน แต่เมื่อใช้ขนาดกลุ่มตัวอย่างที่มากกว่า 1,000 คน และค่าเฉลี่ยความสามารถของทั้งสองกลุ่มแตกต่างกัน และ ข้อสอบที่มีค่าอำนาจจำแนกสูง ค่าความยากค่าวิธี SIB มีอัตราการเกิดความคลาดเคลื่อนแบบที่ 1 ต่ำกว่าวิธี MH

นารายานาน และสวามินาทาน (Narayanan and Swaminathan, 1994 : 315 - 328) ได้ศึกษาเปรียบเทียบผลการตรวจสอบความลำเอียงแบบยูนิฟอร์มด้วยวิธี MH และวิธี SIB โดยใช้ข้อมูลจำลองภายใต้เงื่อนไข 1,296 เงื่อนไข แล้วทำการคำนวณซ้ำ 100 ครั้ง ในแต่ละเงื่อนไขซึ่งประกอบด้วย

1. ขนาดกลุ่มตัวอย่างที่ต่างกัน 3 ขนาด คือ กลุ่มเปรียบเทียบใช้ขนาด 100, 200 และ 300 คน กลุ่มอ้างอิงใช้ขนาด 300, 500 และ 1,000 คน วางไว้กันได้กลุ่มตัวอย่างทั้งหมด 9 กลุ่ม
2. ความแตกต่างในการแจกแจงความสามารถ 3 ระดับ ระดับที่ 1 ให้ทั้งสองกลุ่มมีค่าเฉลี่ยการแจกแจงความสามารถที่ระดับ 0.0 ระดับที่ 2 ให้ค่าเฉลี่ยความสามารถในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 0.0 และ 0.5 ตามลำดับ และระดับที่ 3 ให้ค่าเฉลี่ยความสามารถในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 0.0 และ -1.0 ส่วนเบี่ยงเบนมาตรฐานของทั้ง 3 ระดับเป็น 1.0
3. ร้อยละของข้อสอบที่ลำเอียงมี 2 ขนาดคือ ร้อยละ 10 และ ร้อยละ 20 จากแบบทดสอบจำนวน 40 ข้อ
4. ขนาดของความลำเอียง 4 ระดับ คือ 0.4, 0.6, 0.8 และ 1.0
5. ชนิดของแบบทดสอบ 6 ชนิดคือ ชนิดของแบบทดสอบที่ประกอบด้วยค่าอำนาจจำแนกและค่าความยาก 3 ระดับ คือ สูง ปานกลาง และต่ำ วางไว้กัน แต่คัดแบบทดสอบที่ประกอบด้วยค่าอำนาจจำแนก และค่าความยากสูง ปานกลาง และต่ำ ทั้งสองค่าออก จำนวน 3 ฉบับ

ผลการวิจัยพบว่า วิธี MH และ SIB สามารถตรวจค้นความลำเอียงได้ใกล้เคียงกันโดยตรวจค้นได้ดีในขนาดกลุ่มตัวอย่าง 600 คน (กลุ่มอ้างอิงและกลุ่มเปรียบเทียบกลุ่มละ 300 คน) ขนาดของความลำเอียงทั้ง 4 ขนาด มีผลต่อการตรวจค้นความลำเอียงทั้งสองวิธีอย่างมีนัยสำคัญทางสถิติ ความแตกต่างในการแจกแจงความสามารถไม่มีผลต่อวิธี SIB แต่มีผลต่อวิธี MH ที่เป็นเช่นนี้อาจเป็นเพราะ วิธี SIB ได้มีการปรับแก้การถดถอย ขนาดของแบบทดสอบและร้อยละของข้อสอบที่ลำเอียงไม่มีผลต่อวิธีทั้งสอง สำหรับการเกิดความคลาดเคลื่อนแบบที่ 1 วิธี MH มีอัตราการเกิดอยู่ในระดับปกติ แต่วิธี SIB จะเพิ่มขึ้นเล็กน้อยในกลุ่มที่มีความสามารถเท่ากันแต่ในกลุ่มที่มีความสามารถไม่เท่ากัน การเกิดความคลาดเคลื่อนแบบที่ 1 จะสูงทั้งสองวิธี ซึ่งการเพิ่มความยาวของ

แบบทดสอบจะทำให้ความเชื่อมั่นสูงขึ้น ซึ่งมีผลทำให้การเกิดความคลาดเคลื่อนแบบที่ 1 น่าจะมีอัตราลดลง

จากการศึกษาเอกสารและงานวิจัยดังกล่าวข้างต้น จะเห็นว่า งานวิจัยส่วนใหญ่จะทำการศึกษาเปรียบเทียบวิธีตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีที่แตกต่างกัน โดยแบ่งได้เป็น 2 แนวทางใหญ่ ๆ คือ วิธีตรวจสอบข้อสอบที่ลำเอียงที่ยึดหลักของทฤษฎีการวัดผลแบบดั้งเดิมและวิธีตรวจสอบข้อสอบที่ลำเอียงที่ยึดหลักของทฤษฎีการตอบข้อคำถาม ซึ่งพบว่า วิธีที่ยึดตามแนวของทฤษฎีการตอบข้อคำถามให้ผลการวิเคราะห์ที่ถูกต้องแม่นยำมากกว่าวิธีที่ยึดตามแนวของทฤษฎีการวัดผลแบบดั้งเดิม แต่ไม่เหมาะสมที่จะนำมาใช้ในทางปฏิบัติ เพราะต้องใช้กลุ่มตัวอย่างขนาดใหญ่ การคำนวณและการแปลผลซับซ้อน ยุ่งยากและต้องเสียค่าใช้จ่ายสูง สำหรับตัวแปรความลำเอียงที่นิยมทำการศึกษาเช่น ภาษา เพศ วัฒนธรรม สภาพภูมิศาสตร์ เป็นต้น นอกจากนี้ ยังพบว่า มีตัวแปรบางตัวที่ส่งผลต่อการตรวจสอบความลำเอียงของข้อสอบ เช่น ความยาวของแบบทดสอบ ขนาดกลุ่มผู้สอบ การแจกแจงความสามารถในแต่ละกลุ่มย่อย เป็นต้น ซึ่งในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยมีความสนใจที่จะศึกษาการเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษาที่มีระดับความง่ายของข้อสอบแตกต่างกัน โดยตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล และวิธีชิปเทสต์ ดังนั้น จึงได้กำหนดสมมติฐานแบบไม่มีทิศทางขึ้น

สมมติฐานของการศึกษาค้นคว้า

1. ข้อสอบที่มีความลำเอียง ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน มีจำนวนข้อแตกต่างกัน
2. ข้อสอบที่ความลำเอียง ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน มีจำนวนข้อแตกต่างกัน
3. ความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน มีค่าแตกต่างกัน
4. ความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 25 โรงเรียน 133 ห้องเรียน มีจำนวนนักเรียนทั้งหมด 5,261 คน จำแนกออกเป็น 4 ขนาด คือ โรงเรียนขนาดใหญ่พิเศษ จำนวน 1 โรงเรียน จำนวน 20 ห้องเรียน จำนวนนักเรียน 937 คน โรงเรียนขนาดใหญ่ จำนวน 4 โรงเรียน จำนวน 35 ห้องเรียน จำนวนนักเรียน 1,450 คน โรงเรียนขนาดกลาง จำนวน 13 โรงเรียน จำนวน 64 ห้องเรียน จำนวนนักเรียน 2,468 คน โรงเรียนขนาดเล็ก จำนวน 7 โรงเรียน จำนวน 14 ห้องเรียน จำนวนนักเรียน 406 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี จำนวน 1,041 คน ซึ่งมีความเชื่อมั่นได้ร้อยละ 95 ($\alpha = .05$) (ล้วน สายยศ และอังคณา สายยศ. 2536 : 261 ; อ้างอิงมาจาก Yamane. 1967) โดยแบ่งเป็นกลุ่มตัวอย่างที่ใช้ในการตรวจสอบจำนวนข้อสอบที่มีความลำเอียงและกลุ่มตัวอย่างที่ใช้ในการคำนวณค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้วโดยใช้การสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ซึ่งมีลำดับขั้นตอนในการสุ่ม ดังนี้

ขั้นที่ 1 สุ่มห้องเรียนจากแต่ละขนาดของโรงเรียนด้วยสัดส่วน .20 หรือร้อยละ 20 ด้วยวิธีสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) และมีห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้จำนวนห้องเรียน 26 ห้องเรียน ที่มีจำนวนนักเรียนทั้งหมด 1,041 คน

ขั้นที่ 2 แบ่งห้องเรียนที่สุ่มได้จากขั้นที่ 1 โดยวิธีสุ่มอย่างง่าย (Simple Random Sampling) มีห้องเรียนเป็นหน่วยการสุ่ม โดยแบ่งเป็นกลุ่มตัวอย่างที่ใช้ในการตรวจสอบข้อสอบที่มีความลำเอียง จำนวน 16 ห้องเรียน มีจำนวนนักเรียน 641 คน และกลุ่มตัวอย่างที่ใช้ในการหาค่า

ความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จำนวน 10 ห้องเรียน มีจำนวนนักเรียน 400 คน ดังรายละเอียดที่เสนอในตาราง 3

ตาราง 3 จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามขนาดโรงเรียน

ขนาด	โรงเรียน	จำนวนนักเรียนที่ใช้ในการตรวจสอบข้อสอบที่ลำเอียง(ห้อง)	จำนวนนักเรียนที่ใช้ในการหาค่าความเชื่อมั่นหลังคัดข้อสอบที่ลำเอียงออกแล้ว(ห้อง)
ใหญ่พิเศษ	พระนารายณ์	175(4)	
ใหญ่	พัฒนานิคม	200(5)	
	ชัยบาดาลวิทยา	153(4)	
กลาง	โคกตูมวิทยา		125(4)
	ชัยบาดาลพิทยาคม		120(3)
	ท่าหลวงวิทยาคม	113(3)	
เล็ก	โคกสูงวิทยา		68(2)
	ขุนรามวิทยา		40(1)
รวม		641(16)	400(10)

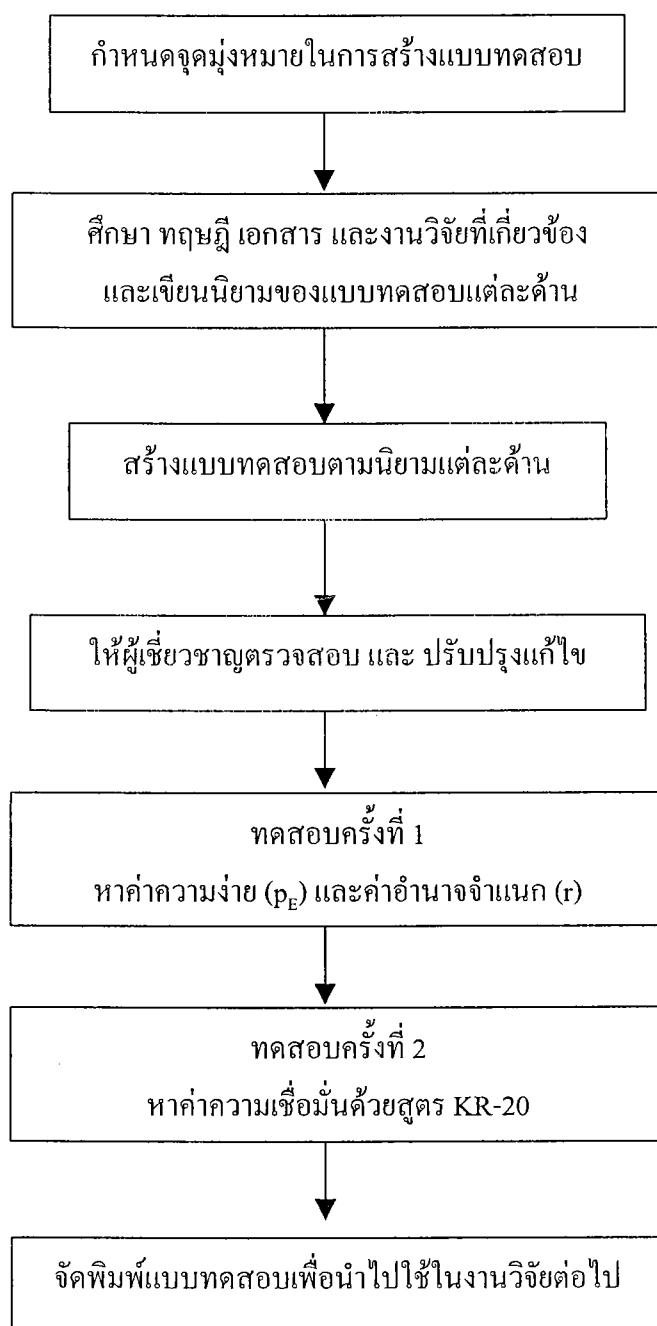
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดความสามารถด้านภาษา ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 60 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง โดยแบ่งออกเป็น 5 ด้าน ดังนี้

- | | |
|---------------------------------|--------------|
| 1. ด้านคำตรงข้าม | จำนวน 10 ข้อ |
| 2. ด้านคำที่มีความหมายใกล้เคียง | จำนวน 10 ข้อ |
| 3. ด้านคำศัพท์สัมพันธ์ | จำนวน 10 ข้อ |
| 4. ด้านความหมายไม่เข้าพวก | จำนวน 10 ข้อ |
| 5. ด้านความเข้าใจภาษา | จำนวน 20 ข้อ |

วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

ในการสร้างเครื่องมือสำหรับการศึกษาค้นคว้า ผู้วิจัยได้ดำเนินการสร้างตามขั้นตอน
ดังนี้



ภาพประกอบ 1 ลำดับขั้นตอนในการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

รายละเอียดขั้นตอนการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

จากแผนภูมิดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า มีรายละเอียด ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อศึกษาเปรียบเทียบผลการตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบและวิธีตรวจสอบความลำเอียงที่ต่างกัน

2. ศึกษาทฤษฎี เอกสาร และงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับวิธีการสร้างแบบทดสอบวัดความสามารถด้านภาษา และเขียนนิยามความหมายของแบบทดสอบทั้ง 5 ด้าน ได้แก่ ด้านคำตรงข้าม ด้านคำที่มีความหมายใกล้เคียง ด้านคำศัพท์สัมพันธ์ ด้านความหมายไม่เข้าพวก และด้านความเข้าใจภาษา

3. สร้างแบบทดสอบวัดความสามารถด้านภาษาแต่ละด้านตามคำนิยามในข้อ 2 โดยสร้างเป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก ทั้งหมด 220 ข้อ มีรายละเอียดดังนี้

3.1 ด้านคำตรงข้าม	จำนวน	40	ข้อ
3.2 ด้านคำที่มีความหมายใกล้เคียง	จำนวน	40	ข้อ
3.3 ด้านคำศัพท์สัมพันธ์	จำนวน	40	ข้อ
3.4 ด้านความหมายไม่เข้าพวก	จำนวน	40	ข้อ
3.5 ด้านความเข้าใจภาษา	จำนวน	60	ข้อ

4. นำแบบทดสอบวัดความสามารถด้านภาษา จำนวน 220 ข้อ ที่สร้างขึ้นจากข้อ 3 ไปให้ผู้เชี่ยวชาญด้านการวัดผลการศึกษา จำนวน 5 ท่าน ตรวจสอบความเที่ยงตรงเชิงพินิจ (Face Validity) โดยพิจารณาความสอดคล้องระหว่างข้อคำถามกับคุณลักษณะที่นิยามไว้ ซึ่งข้อสอบที่เหมาะสมควรมีค่า IOC ตั้งแต่ 0.5 ขึ้นไป ผลปรากฏว่า มีข้อคำถามอยู่จำนวน 55 ข้อ ที่ต้องปรับปรุงและแก้ไข ซึ่งผู้วิจัยได้ดำเนินการปรับปรุงและแก้ไขข้อบกพร่องต่าง ๆ ตามที่ผู้เชี่ยวชาญเสนอแนะจนได้แบบทดสอบวัดความสามารถด้านภาษาที่สมบูรณ์ จำนวน 220 ข้อ

5. นำแบบทดสอบวัดความสามารถด้านภาษาที่ปรับปรุงแก้ไขแล้วจาก ข้อ 4 ไปทดลองสอบครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนโคกสำโรงวิทยา จำนวน 200 คน แล้วนำผลการสอบมาวิเคราะห์หาค่าความง่าย (p_E) และ ค่าอำนาจจำแนก (r) ของแบบทดสอบเป็นรายข้อ จากนั้นคัดเลือกข้อสอบที่มีค่าความง่าย (p_E) 3 ระดับ คือ ข้อสอบที่มีระดับความง่ายต่ำ มีค่า p_E ตั้งแต่ 0.20 ถึง 0.39 จำนวน 20 ข้อ ข้อสอบที่มีระดับความง่ายปานกลาง มีค่า p_E ตั้งแต่ 0.40 ถึง 0.59 จำนวน 20 ข้อ และข้อสอบที่มีระดับความง่ายสูง มีค่า p_E ตั้งแต่ 0.60 ถึง 0.80 จำนวน 20 ข้อ

รวมคัดเลือกข้อสอบไว้ทั้งหมด 60 ข้อ โดยข้อสอบแต่ละข้อจะต้องมีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไป และครอบคลุมเนื้อหาในแบบทดสอบวัดความสามารถด้านภาษาทั้ง 5 ด้าน ๆ ละ อย่างน้อย 2 ข้อ ผลการวิเคราะห์ข้อสอบมีดังนี้ (รายละเอียด แสดงในภาคผนวก ข)

5.1 ด้านคำตรงข้าม มีจำนวนข้อสอบทั้งหมด 40 ข้อ มีค่า p_E ตั้งแต่ 0.21 ถึง 0.84 และมีค่า r ตั้งแต่ 0.11 ถึง 0.66 เมื่อคัดเลือกข้อสอบที่อยู่ในเกณฑ์ไว้จำนวน 10 ข้อ มีค่า p_E ตั้งแต่ 0.21 ถึง 0.69 และมีค่า r ตั้งแต่ 0.23 ถึง 0.65

5.2 ด้านคำที่มีความหมายใกล้เคียง มีจำนวนข้อสอบทั้งหมด 40 ข้อ มีค่า p_E ตั้งแต่ 0.23 ถึง 0.72 และมีค่า r ตั้งแต่ -0.11 ถึง 0.59 เมื่อคัดเลือกข้อสอบที่อยู่ในเกณฑ์ไว้จำนวน 10 ข้อ มีค่า p_E ตั้งแต่ 0.35 ถึง 0.71 และมีค่า r ตั้งแต่ 0.24 ถึง 0.54

5.3 ด้านคำศัพท์สัมพันธ์ มีจำนวนข้อสอบทั้งหมด 40 ข้อ มีค่า p_E ตั้งแต่ 0.19 ถึง 0.70 และมีค่า r ตั้งแต่ 0.13 ถึง 0.80 เมื่อคัดเลือกข้อสอบที่อยู่ในเกณฑ์ไว้จำนวน 10 ข้อ มีค่า p_E ตั้งแต่ 0.34 ถึง 0.62 และมีค่า r ตั้งแต่ 0.22 ถึง 0.80

5.4 ด้านความหมายไม่เข้าพวก มีจำนวนข้อสอบทั้งหมด 40 ข้อ มีค่า p_E ตั้งแต่ 0.16 ถึง 0.84 และมีค่า r ตั้งแต่ -0.22 ถึง 0.70 เมื่อคัดเลือกข้อสอบที่อยู่ในเกณฑ์ไว้จำนวน 10 ข้อ มีค่า p_E ตั้งแต่ 0.31 ถึง 0.67 และมีค่า r ตั้งแต่ 0.26 ถึง 0.70

5.5 ด้านความเข้าใจภาษา มีจำนวนข้อสอบทั้งหมด 60 ข้อ มีค่า p_E ตั้งแต่ 0.18 ถึง 0.78 และมีค่า r ตั้งแต่ 0.18 ถึง 0.72 เมื่อคัดเลือกข้อสอบที่อยู่ในเกณฑ์ไว้จำนวน 20 ข้อ มีค่า p_E ตั้งแต่ 0.24 ถึง 0.69 และมีค่า r ตั้งแต่ 0.20 ถึง 0.65

6. นำแบบทดสอบวัดความสามารถด้านภาษาที่คัดเลือกไว้ทั้งหมด จำนวน 60 ข้อ ไปทดลองสอบครั้งที่ 2 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนบ้านหมี่วิทยา จำนวน 200 คน จากนั้นนำผลการสอบมาวิเคราะห์หาค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ ด้วยวิธีของคูเดอร์ริชาร์ดสัน (KR-20) ผลปรากฏว่า ได้ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับเท่ากับ 0.78

7. จัดพิมพ์แบบทดสอบเพื่อนำไปใช้ในการวิจัยต่อไป

ตัวอย่างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

แบบทดสอบวัดความสามารถด้านภาษา

1. ด้านคำตรงข้าม

คำชี้แจง ให้นักเรียนพิจารณาคำหรือข้อความ ที่มีความหมายตรงข้ามกับคำที่ขีดเส้นใต้ในแต่ละข้อ

(0) คิดถึง

ก. ลืมเลือน

ข. จางหาย

ค. ขาดสติ

ง. ไม่ใส่ใจ

จ. เกล็ดขัง

เฉลย ข้อ ก

2. ด้านคำที่มีความหมายใกล้เคียง

คำชี้แจง ให้นักเรียนพิจารณาคำหรือข้อความ ที่มีความหมายใกล้เคียงมากที่สุดกับคำที่ขีดเส้นใต้ในแต่ละข้อ

(0) อาภัพ

ก. ต่ำต้อย

ข. ผิดหวัง

ค. ล้มเหลว

ง. อับโชค

จ. โศกเศร้า

เฉลย ข้อ ง

3. ด้านคำศัพท์สัมพันธ์

คำชี้แจง ให้นักเรียนพิจารณาคำหรือข้อความ ที่มีความสัมพันธ์หรือเกี่ยวข้องมากที่สุด กับคำที่กำหนดให้ในแต่ละข้อ

(0) ลิง

- ก. นารัก
- ข. ชุกชน
- ค. ร่าเริง
- ง. ว่องไว
- จ. นิลาค

เฉลย ข้อ ข

4. ด้านความหมายไม่เข้าพวก

คำชี้แจง ให้นักเรียนพิจารณาว่าในแต่ละข้อต่อไปนี้ ข้อใดมีความหมายแตกต่าง จากข้ออื่น ๆ

- (0) ก. ไปให้ไกล ๆ
- ข. บอกว่าย่ำมายุ่ง
 - ค. อย่ากังวลใจได้ไหม
 - ง. อยากอยู่เงียบ ๆ
 - จ. ปรีक्षाอะไรหน่อย

เฉลย ข้อ จ

5. ด้านความเข้าใจภาษา

คำชี้แจง อ่านบทประพันธ์ต่อไปนี้ แล้วตอบคำถามข้อ (0) – (00)

“ทั้งนี้เป็นต้นด้วยผลเหตุ	จะอาเพศเพราะผู้เป็นใหญ่
มิได้พิจารณาข้าไท	เคยใช้ก็เลี้ยงด้วยเมตตา
ไม่รู้รอบประกอบในราชกิจ	ประพบัติการแต่ผัดด้วยอิงฉา
สุภายิตท่านกล่าวเป็นรวมา	จะแต่งตั้งเสนาบดี
ไม่ควรอย่าให้อัครฐาน	จะเสียการแผ่นดินกรุงศรี
เพราะไม่เชื่อดำนานโบราณคดี	จึงเสียทิวังศ์กษัตรา”

(0) คำประพันธ์ข้างต้น กล่าวถึงเรื่องใดเป็นสำคัญ

- ก. การปกครองบ้านเมือง
- ข. การทำศึกสงคราม
- ค. ความเมตตากรุณา
- ง. การแต่งตั้งข้าราชการ
- จ. การรักษาเอกราช

เฉลย ข้อ ง

(00) *เสนาบดี* เปรียบได้กับตำแหน่งใดในปัจจุบัน

- ก. นักการเมือง
- ข. นายกรัฐมนตรี
- ค. สมาชิกวุฒิสภา
- ง. สมาชิกสภาผู้แทนราษฎร
- จ. รัฐมนตรีว่าการกระทรวง

เฉลย ข้อ จ

การตรวจให้คะแนน

การตรวจให้คะแนนแบบทดสอบวัดความสามารถด้านภาษา ทั้ง 5 ด้าน ใช้เกณฑ์การตรวจให้คะแนน คือ ตอบถูก ให้ 1 คะแนน และ ตอบผิดหรือไม่ตอบให้ 0 คะแนน

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. นำหนังสือจากบัณฑิตวิทยาลัยไปติดต่อขอความร่วมมือกับโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนด วัน เวลา ในการดำเนินการสอบ
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่จะทำการสอบในแต่ละครั้ง
3. วางแผนดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการสอบด้วยตนเอง
4. ชี้แจงให้นักเรียนที่เป็นกลุ่มตัวอย่างทราบวัตถุประสงค์ของการสอบ และขอความร่วมมือในการสอบ เพื่อให้ได้ผลสอบตามความเป็นจริงมากที่สุด
5. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัย กลุ่มที่ 1 จำนวน 641 คน ในช่วงเดือนมิถุนายน 2544 เพื่อนำผลการสอบมาศึกษาผลการตรวจสอบความลำเอียงของข้อสอบในด้านจำนวนข้อสอบที่มีความลำเอียง
6. นำแบบทดสอบที่คัดข้อสอบที่มีความลำเอียงออกแล้ว จากการตรวจสอบความลำเอียงของข้อสอบแต่ละวิธี ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัยกลุ่มที่ 2 จำนวน 400 คน ในช่วงเดือนมิถุนายน ถึงเดือนกรกฎาคม 2544 เพื่อนำผลการสอบมาวิเคราะห์หาค่าความเชื่อมั่นของแบบทดสอบ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการหาคุณภาพเครื่องมือ
 - 1.1 หาค่าความง่าย (p_E) และค่าอำนาจจำแนก (r) โดยเปิดจากตารางสำเร็จของ จุง-เตห์-ฟาน (Chung-Teh-Fan) โดยใช้เทคนิค 27% ในการแบ่งกลุ่มสูงกลุ่มต่ำ (ล้วน สายยศ และ อังคณา สายยศ. 2539 : 191)
 - 1.2 หาค่าความเชื่อมั่นของแบบทดสอบ ก่อนและหลังการคัดเลือกข้อสอบที่มีความลำเอียงออก ด้วยวิธีของคูเดอร์ ริชาร์ดสัน (KR-20) (ล้วน สายยศ และ อังคณา สายยศ. 2536 : 168)

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

เมื่อ	r_{tt}	แทน ความเชื่อมั่นของแบบทดสอบทั้งฉบับ
	n	แทน จำนวนข้อสอบของแบบทดสอบ

p	แทน สัดส่วนของผู้ตอบถูกในข้อหนึ่ง ๆ
q	แทน สัดส่วนของผู้ตอบผิดในข้อหนึ่ง ๆ ($1 - p$)
s_i^2	แทน ความแปรปรวนของแบบทดสอบทั้งฉบับ

2. สถิติที่ใช้ในการตรวจสอบข้อสอบที่ลำเอียง

2.1 วิธีชิปเทสท์ (SIBTEST) โดยใช้สูตรดังนี้ (Narayanan and Swainathan. 1994 : 317)

$$\hat{\beta}_U = \sum_{k=0}^n \hat{p}_k (\bar{y}_{Rk} - \bar{y}_{Fk})$$

เมื่อ	$\hat{p}_k$	แทน	สัดส่วนของผู้สอบกลุ่มเปรียบเทียบที่ตอบแบบทดสอบ ย่อยที่ไม่ ลำเอียงได้ถูก k ข้อ
	$\bar{y}_{Rk}$	แทน	คะแนนเฉลี่ยของผู้สอบทุกคนในกลุ่มอ้างอิงที่ตอบแบบ ทดสอบ ย่อยที่ศึกษาได้ถูก k ข้อ
	$\bar{y}_{Fk}$	แทน	คะแนนเฉลี่ยของผู้สอบทุกคนในกลุ่มเปรียบเทียบที่ตอบ แบบทดสอบย่อยที่ศึกษาได้ถูก k ข้อ

2.1.1 สถิติที่ใช้ในการทดสอบค่า SIBTEST เมื่อค่า B_U หรือ ค่า Z_{SIB} ที่คำนวณ
จากสูตรแตกต่างจากค่า Z_α ที่ $N(0,1)$ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

$$B_U = \hat{\beta}_U / \hat{\sigma}(\hat{\beta}_U)$$

$$\hat{\sigma}(\hat{\beta}_U) = \left\{ \sum_{k=0}^n \hat{p}_k^2 \left[\frac{1}{j_{Rk}} \hat{\sigma}^2(y/R, k) + \frac{1}{j_{Fk}} \hat{\sigma}^2(y/F, k) \right] \right\}^2$$

เมื่อ	$\hat{\sigma}(\hat{\beta}_U)$	แทน	ค่าประมาณความคลาดเคลื่อนมาตรฐานของ B_U
	j_{Rk}	แทน	จำนวนผู้สอบในกลุ่มอ้างอิงในช่วงคะแนน k
	j_{Fk}	แทน	จำนวนผู้สอบในกลุ่มเปรียบเทียบในช่วงคะแนน k

โดยที่

กลุ่มอ้างอิง แทน กลุ่มที่ได้เปรียบในการตอบข้อสอบได้แก่ กลุ่มผู้สอบเพศชาย
กลุ่มเปรียบเทียบ แทน กลุ่มที่เสียเปรียบในการตอบข้อสอบได้แก่ กลุ่มผู้สอบเพศหญิง

2.2 วิธีของแมนเทล-แฮนส์เซล (Mantel-Haenszel) โดยใช้สูตรดังนี้ (Narayanan and Swainathan. 1994 : 316)

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

เมื่อ	A_j	แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบถูก
	B_j	แทน จำนวนผู้สอบกลุ่มอ้างอิงที่ตอบข้อสอบผิด
	C_j	แทน จำนวนผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบถูก
	D_j	แทน จำนวนผู้สอบกลุ่มเปรียบเทียบที่ตอบข้อสอบผิด
	N_j	แทน จำนวนผู้สอบทั้งสองกลุ่มที่ได้รับคะแนนรวมจากการทดสอบในช่วงคะแนน j
	α_{MH}	แทน ค่าดัชนีความลำเอียงของข้อสอบ

2.2.1 สถิติที่ใช้ทดสอบค่าความแตกต่างจาก 1.00 อย่างมีนัยสำคัญที่ $df = 1$ ใช้สูตรดังนี้

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

เมื่อ	χ_{MH}^2	แทน ค่าการทดสอบความลำเอียงอย่างมีนัยสำคัญของข้อสอบ
	$E(A_j)$ และ $Var(A_j)$	แทน ค่าคาดหวังและค่าความแปรปรวนของ A ที่ระดับคะแนน j

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

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

3. สถิติที่ใช้ในการทดสอบความแตกต่างของจำนวนข้อสอบที่มีความลำเอียงจากการตรวจสอบความลำเอียงด้วยวิธีและระดับความง่ายของข้อสอบต่างกัน โดยใช้สูตรดังนี้ (ล้วน สายยศ และ อังคณา สายยศ. 2536 : 93)

$$Z = \frac{p_1 - p_2}{S_{p1-p2}}$$

เมื่อ Z แทน ค่าการแจกแจงของ Z

p_1 แทน สัดส่วนของข้อสอบที่ลำเอียงของกลุ่มที่ 1 คำนวณจากจำนวนข้อสอบที่ลำเอียงของกลุ่มที่ 1 หารด้วยข้อสอบทั้งหมด

$$p_1 = \frac{f_1}{n_1}$$

p_2 แทน สัดส่วนของข้อสอบที่ลำเอียงของกลุ่มที่ 2 คำนวณจากจำนวนข้อสอบที่ลำเอียงของกลุ่มที่ 2 หารด้วยข้อสอบทั้งหมด

$$p_2 = \frac{f_2}{n_2}$$

S_{p1-p2} แทน ค่าความคลาดเคลื่อนมาตรฐานของความแตกต่างระหว่างสัดส่วนที่วิเคราะห์ความลำเอียงของ 2 กลุ่ม ซึ่งมีสูตรดังนี้

$$S_{p1-p2} = \sqrt{pq \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

เมื่อ p แทน สัดส่วนของข้อสอบที่มีความลำเอียงที่ตรวจสอบความลำเอียงด้วยวิธีตรวจสอบทั้งสองวิธี ซึ่งคำนวณมาจากจำนวนข้อสอบที่ลำเอียงที่ตรวจสอบด้วยวิธีที่ 1 และวิธีที่ 2 หารด้วยจำนวนข้อสอบของทั้ง 2 วิธี

$$p = \frac{f_1 + f_2}{n_1 + n_2} \quad \text{และ} \quad q = 1 - p$$

4. สถิติที่ใช้ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกรหัสสอบที่ลำเอียงออกแล้ว จากการตรวจสอบความลำเอียงด้วยวิธีและระดับค่าความง่ายของข้อสอบที่ต่างกัน โดยทดสอบด้วยไคสแควร์ (χ^2) (Snedecor and Cochran. 1967 : 187)

$$\chi^2 = \sum (n_i - 3) Z_i^2 - \frac{[\sum (n_i - 3) Z_i]^2}{\sum (n_i - 3)} ; df = k-1$$

เมื่อ	χ^2	แทน	ค่าไคสแควร์
	Z_i	แทน	คะแนนมาตรฐานซึ่งแปลงจากสัมประสิทธิ์ความเชื่อมั่นด้วยตาราง Fisher's Z
	n_i	แทน	จำนวนนักเรียนในแต่ละกลุ่ม
	df	แทน	degree of freedom
	k	แทน	จำนวนข้อสอบในแบบทดสอบ

5. ถ้าพบว่าความเชื่อมั่นที่ทดสอบมีความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ตรวจสอบความแตกต่างเป็นรายคู่ โดยใช้สูตรทดสอบความแตกต่างของคะแนนมาตรฐาน (Z) (Ferguson. 1981 : 196)

$$Z = \frac{Z_{R_1} - Z_{R_2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

เมื่อ	Z	แทน	คะแนนมาตรฐานของไค์งปกติ
	Z_{R_1}, Z_{R_2}	แทน	คะแนนมาตรฐานที่แปลงจากค่าสัมประสิทธิ์ความเชื่อมั่นด้วยตาราง Fisher's Z
	N_1, N_2	แทน	จำนวนนักเรียนในกลุ่มย่อย 2 กลุ่ม ตามลำดับ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการเสนอผลการวิเคราะห์ข้อมูล เพื่อความสะดวกและเข้าใจตรงกัน ผู้วิจัยจึงได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

P_E	แทน	ค่าความง่ายของแบบทดสอบ
r	แทน	ค่าอำนาจจำแนกของแบบทดสอบ
k	แทน	จำนวนข้อของแบบทดสอบ
k_b	แทน	จำนวนข้อสอบที่มีความลำเอียง
β_u	แทน	ค่าดัชนีความลำเอียงของข้อสอบที่ได้จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสท์
β_b	แทน	ค่าดัชนีความลำเอียงของข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ที่ได้จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสท์
α_{MH}	แทน	ค่าดัชนีความลำเอียงของข้อสอบที่ได้จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล
α_b	แทน	ค่าดัชนีความลำเอียงของข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ที่ได้จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล
SIB	แทน	การตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสท์
MH	แทน	การตรวจสอบความลำเอียงของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล
r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบที่คำนวณด้วยวิธีของคูเคอร์ ริชาร์ดสัน (KR 20)
χ^2	แทน	ค่าไคส์แควร์
Z	แทน	ค่านัยสำคัญของการทดสอบด้วยสถิติ Z-test
Z_r	แทน	คะแนนมาตรฐานจากค่าสัมประสิทธิ์ของความเชื่อมั่นด้วยตาราง Fisher-Z

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังต่อไปนี้

1. การตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน
2. การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน
3. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

ผลการวิเคราะห์ข้อมูล

1. การตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

ผู้วิจัยได้นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่างจำนวน 641 คน แบ่งเป็นกลุ่มตัวอย่างเพศชาย ซึ่งเป็นกลุ่มอ้างอิงหรือกลุ่มที่ได้เปรียบในการทำแบบทดสอบ จำนวน 316 คน กลุ่มตัวอย่างเพศหญิง ซึ่งเป็นกลุ่มเปรียบเทียบกับหรือกลุ่มที่เสียเปรียบในการทำแบบทดสอบ จำนวน 325 คน จากนั้นจึงนำผลการตอบแบบทดสอบมาตรวจสอบความลำเอียงของข้อสอบ ด้วยวิธีตรวจสอบความลำเอียง 2 วิธี คือ วิธีชิปเทสท์ และ วิธีแมนเทล-แฮนส์เซล โดยจำแนกตามระดับความง่ายของข้อสอบ 3 ระดับ คือ กลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง ได้ผลการวิเคราะห์ดังตาราง 4 ดังนี้

ตาราง 4 การตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

ระดับความ ง่ายของ ข้อสอบ	k	วิธี SIB				วิธี MH			
		β_u	k_b	β_b	ข้อสอบที่ ลำเอียง	α_{MH}	k_b	α_b	ข้อสอบที่ ลำเอียง
ต่ำ	20	-0.084 ถึง 0.136	4	-0.078 ถึง 0.136	3, 21, 29 และ 45	0.64 ถึง 2.22	4	0.64 ถึง 2.22	3,21,29 และ 45
ปานกลาง	20	-0.112 ถึง 0.331	3	-0.112 ถึง 0.331	23, 38 และ 49	0.55 ถึง 5.85	3	0.55 ถึง 5.85	23, 38 และ 49
สูง	20	-0.079 ถึง 0.058	0	-	-	0.60 ถึง 1.67	1	1.67	1

จากตาราง 4 พบว่า ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ ซึ่งเป็นข้อสอบที่ค่อนข้างยาก และมีจำนวนข้อสอบทั้งหมด 20 ข้อ เมื่อตรวจสอบความลำเอียงด้วยวิธีชิปเทสต์ จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง -0.084 ถึง 0.136 โดยมีข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 4 ข้อ ได้แก่ ข้อสอบข้อที่ 3,21,29 และ 45 เมื่อตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง 0.64 ถึง 2.22 โดยมีข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 4 ข้อ ได้แก่ ข้อสอบข้อที่ 3,21,29 และ 45 ดังจะเห็นได้ว่า ทั้งวิธีชิปเทสต์และวิธีแมนเทิล-แฮนส์เซล สามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ตรงกันทุกข้อ

ส่วนในกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง ซึ่งมีจำนวนข้อสอบทั้งหมด 20 ข้อ เมื่อตรวจสอบความลำเอียงด้วยวิธีชิปเทสต์ จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง -0.112 ถึง 0.331 โดยมีข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 3 ข้อ ได้แก่ ข้อสอบข้อที่ 23,38 และ 49 เมื่อตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง 0.55 ถึง 5.85 โดยมีข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 3 ข้อ ได้แก่ ข้อสอบข้อที่ 23,38 และ 49 ดังจะเห็นได้ว่า ทั้งวิธี

ชิปเทสต์และวิธีแมนเทิล-แฮนส์เซล ก็สามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ตรงกันทุกข้อ

สำหรับในกลุ่มข้อสอบที่มีระดับความง่ายสูง ซึ่งเป็นข้อสอบที่ค่อนข้างง่าย และมีจำนวนข้อสอบทั้งหมด 20 ข้อ เมื่อตรวจสอบความลำเอียงด้วยวิธีชิปเทสต์ จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง -0.079 ถึง 0.058 โดยเป็นข้อสอบที่มีความลำเอียงอย่างไม่มีนัยสำคัญทางสถิติทุกข้อ แสดงว่า วิธีชิปเทสต์จะไม่ตรวจพบข้อสอบที่มีความลำเอียงในข้อสอบที่ค่อนข้างง่าย และเมื่อตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า จะมีค่าดัชนีความลำเอียงของข้อสอบอยู่ระหว่าง 0.60 ถึง 1.67 โดยมีข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ $.05$ อยู่เพียง 1 ข้อ ได้แก่ ข้อสอบข้อที่ 1

เมื่อพิจารณาผลการตรวจสอบข้อสอบที่ลำเอียงด้วยวิธีเดียวกัน พบว่า วิธีชิปเทสต์จะตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้มากที่สุด ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ คือ จำนวน 4 ข้อ รองลงมาคือ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง จำนวน 3 ข้อ แต่ไม่สามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงในกลุ่มข้อสอบที่มีระดับความง่ายสูง ส่วนวิธีแมนเทิล-แฮนส์เซลจะตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้มากที่สุด ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ คือ จำนวน 4 ข้อ รองลงมา คือ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง จำนวน 3 ข้อ และตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้น้อยที่สุดในกลุ่มข้อสอบที่มีระดับความง่ายสูง คือ จำนวน 1 ข้อ

2. การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

จากผลการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซล จำแนกตามระดับความง่ายของข้อสอบ พบว่า สามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ใกล้เคียงกันในแต่ละวิธีตรวจสอบและในแต่ละระดับความง่ายของข้อสอบ ดังนั้น เพื่อให้เกิดความชัดเจนยิ่งขึ้น ผู้วิจัยจึงทำการเปรียบเทียบความแตกต่างของจำนวนข้อสอบที่มีความลำเอียงโดยใช้สถิติ Z-test เพื่อทดสอบความแตกต่างอย่างมีนัยสำคัญทางสถิติของจำนวนข้อสอบที่มีความลำเอียง ดังแสดงในตาราง 5 ถึง ตาราง 6 ดังนี้

ตาราง 5 การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างระดับความง่ายของข้อสอบ
ต่างกัน โดยใช้การตรวจสอบความลำเอียงวิธีเดียวกัน

วิธีตรวจสอบความลำเอียง ระดับความง่ายของข้อสอบ(จำนวนข้อ)	ระดับความง่ายของข้อสอบ (จำนวนข้อ)		
SIB	ต่ำ (4)	ปานกลาง (3)	สูง (0)
ต่ำ (4)	-	0.416	2.108*
ปานกลาง (3)	-	-	1.801
สูง (0)	-	-	-
MH	ต่ำ (4)	ปานกลาง (3)	สูง (1)
ต่ำ (4)	-	0.416	1.434
ปานกลาง (3)	-	-	1.054
สูง (1)	-	-	-

* มีนัยสำคัญทางสถิติที่ระดับ .05

$$Z_{.05} = 1.96$$

จากตาราง 5 เมื่อตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสท์ พบว่า ข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายปานกลางกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

และเมื่อตรวจสอบความลำเอียงของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า ข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งแสดงว่า ในกลุ่มข้อสอบที่มีความง่ายต่างกันทั้ง 3 ระดับ จะพบจำนวนข้อสอบที่มีความลำเอียงใกล้เคียงกัน

ตาราง 6 การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน

ระดับความง่ายของข้อสอบ	วิธีตรวจสอบความลำเอียง		Z
	SIB	MH	
ต่ำ	4	4	0.000
ปานกลาง	3	3	0.000
สูง	0	1	1.013

$$Z_{.05} = 1.96$$

จากตาราง 6 พบว่า ข้อสอบที่มีความลำเอียงในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง ระหว่างที่ตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซล มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

หลังจากทำการตรวจสอบความลำเอียงของข้อสอบโดยวิธีตรวจสอบความลำเอียงทั้งสองวิธีในแต่ละระดับความง่ายของข้อสอบแล้ว ผู้วิจัยได้นำแบบทดสอบที่คัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 400 คน เพื่อคำนวณหาค่าความเชื่อมั่นของแบบทดสอบ และทำการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ด้วยสถิติไคสแควร์ ได้ผลการวิเคราะห์ดังตาราง 7 ถึง ตาราง 8 ดังนี้

ตาราง 7 ค่าความเชื่อมั่นของแบบทดสอบ ค่า Fisher-Z และค่าการทดสอบนัยสำคัญของค่าความ
เชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างระดับ
ความง่ายของข้อสอบต่างกัน โดยใช้การตรวจสอบความลำเอียงวิธีเดียวกัน

วิธีตรวจสอบความลำเอียง ระดับความง่ายของข้อสอบ	k	r_u	Z_r	χ^2
<u>SIB</u>				
ต่ำ	16	0.606	0.701	1.030
ปานกลาง	17	0.645	0.767	
สูง	20	0.608	0.709	
<u>MH</u>				
ต่ำ	16	0.606	0.701	0.886
ปานกลาง	17	0.645	0.767	
สูง	19	0.622	0.725	

$$\chi^2_{0.5,2} = 5.991$$

จากตาราง 7 พบว่า ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความ
ลำเอียงออกแล้ว เมื่อตรวจสอบความลำเอียงด้วยวิธีชิปเทสต์ มีค่าระหว่าง 0.608 ถึง 0.645 โดยกลุ่ม
ข้อสอบที่มีระดับความง่ายปานกลาง มีค่าความเชื่อมั่นสูงสุดเป็น 0.645 รองลงมาคือ กลุ่มข้อสอบที่
มีระดับความง่ายสูง มีค่าความเชื่อมั่นเป็น 0.608 ส่วนกลุ่มข้อสอบที่มีระดับความง่ายต่ำ มีค่าความ
เชื่อมั่นต่ำที่สุด เท่ากับ 0.606 และเมื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบด้วยสถิติ
ไคสแควร์ พบว่า มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สำหรับค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออก
แล้ว เมื่อตรวจสอบความลำเอียงด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า มีค่าระหว่าง 0.606 ถึง 0.645
โดยกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง มีค่าความเชื่อมั่นสูงสุดเป็น 0.645 รองลงมาคือ กลุ่ม
ข้อสอบที่มีระดับความง่ายสูง มีค่าความเชื่อมั่นเป็น 0.622 ส่วนกลุ่มข้อสอบที่มีระดับความง่ายต่ำ
มีค่าความเชื่อมั่นต่ำที่สุด เท่ากับ 0.606 และเมื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบด้วย
สถิติไคสแควร์ พบว่า มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 8 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน

ระดับความง่ายของข้อสอบ	วิธีตรวจสอบความลำเอียง		Z
	SIB	MH	
ต่ำ	0.606	0.606	0.000
ปานกลาง	0.645	0.645	0.000
สูง	0.608	0.622	0.226

$$Z_{.05} = 1.96$$

จากตาราง 8 พบว่า ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง ระหว่างการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์ และวิธีแมนเทิล-แฮนส์เซด มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุป อภิปรายผล และ ข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน
2. เพื่อเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน
3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างระดับความง่ายของข้อสอบต่างกัน เมื่อตรวจสอบความลำเอียงด้วยวิธีเดียวกัน
4. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบความลำเอียงด้วยวิธีที่ต่างกัน เมื่อระดับความง่ายของข้อสอบเหมือนกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดลพบุรี ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) และห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) จำนวน 26 ห้องเรียน มีจำนวนนักเรียนทั้งสิ้น 1,041 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดความสามารถด้านภาษา ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 60 ข้อ ที่ผู้วิจัยสร้างขึ้นเอง โดยแบ่งออกเป็น 5 ด้าน ดังนี้

- | | |
|---------------------------------|--------------|
| 1. ด้านคำตรงข้าม | จำนวน 10 ข้อ |
| 2. ด้านคำที่มีความหมายใกล้เคียง | จำนวน 10 ข้อ |
| 3. ด้านคำศัพท์สัมพันธ์ | จำนวน 10 ข้อ |

- | | |
|---------------------------|--------------|
| 4. ด้านความหมายไม่เข้าพวก | จำนวน 10 ข้อ |
| 5. ด้านความเข้าใจภาษา | จำนวน 20 ข้อ |

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. นำหนังสือจากบัณฑิตวิทยาลัยไปติดต่อขอความร่วมมือกับโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนด วัน เวลา ในการดำเนินการสอบ
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่จะทำการสอบในแต่ละครั้ง
3. วางแผนดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการสอบด้วยตนเอง
4. ชี้แจงให้นักเรียนที่เป็นกลุ่มตัวอย่างทราบวัตถุประสงค์ของการสอบ และขอความร่วมมือในการสอบ เพื่อให้ได้ผลสอบตามความเป็นจริงมากที่สุด
5. นำแบบทดสอบ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัย กลุ่มที่ 1 จำนวน 641 คน เพื่อนำผลการสอบมาศึกษาผลการตรวจสอบความลำเอียงของข้อสอบในด้านจำนวนข้อสอบที่มีความลำเอียง
6. นำแบบทดสอบที่คัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ด้วยวิธีตรวจสอบความลำเอียงแต่ละวิธี ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัยกลุ่มที่ 2 จำนวน 400 คน เพื่อนำผลการสอบมาวิเคราะห์หาค่าความเชื่อมั่นของแบบทดสอบ

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังต่อไปนี้

1. การตรวจสอบข้อสอบที่ลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน
2. การเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน
3. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จำแนกตามระดับความง่ายของข้อสอบ และวิธีตรวจสอบความลำเอียงที่ต่างกัน

สรุปผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลสรุปได้ดังนี้

1. จำนวนข้อสอบที่มีความลำเอียง เมื่อตรวจสอบด้วยวิธีชิปเทสท์ ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และระหว่างกลุ่มข้อสอบที่มีระดับความง่ายปานกลางกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เมื่อตรวจสอบด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า ข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. จำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง เมื่อตรวจสอบด้วยวิธีชิปเทสท์ และวิธีแมนเทิล-แฮนส์เซล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. ค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

1. จำนวนข้อสอบที่มีความลำเอียง เมื่อตรวจสอบด้วยวิธีชิปเทสท์ ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งผลที่ได้สอดคล้องกับสมมติฐานของการวิจัย และยังสอดคล้องกับงานวิจัยของ ขศอารีย์ รวยธนพานิช (2534:56) ที่ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบปรนัย ชนิดเลือกตอบที่มีช่วงความยากและความยาวของแบบทดสอบต่างกัน ผลการวิจัยพบว่า ข้อสอบที่มี

ช่วงความยากต่างจาก 0.5 ไปยิ่งมากเท่าใด โอกาสที่ผู้สอบจะตอบผิดหรือถูกจะมีมากขึ้นเท่านั้น ซึ่งชวาล แพริตกุล. (2530:9-11) ได้กล่าวถึงค่าความยากของข้อสอบไว้ว่า ข้อสอบที่ง่ายเกินไป ผู้สอบทั้งหมดที่เข้าสอบอาจตอบถูกทั้งหมดหรือเกือบหมด ส่วนข้อสอบที่ยากมาก ผู้สอบอาจตอบผิดทั้งหมดหรือเกือบหมด และ อนันต์ ศรีโสภณ. (2525:166) ยังกล่าวว่า โดยทั่วไป ข้อสอบที่ค่อนข้างง่ายจะมีความคลาดเคลื่อนในการวัดน้อยกว่าข้อสอบที่ค่อนข้างยาก ส่วนข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และระหว่างกลุ่มข้อสอบที่มีระดับความง่ายปานกลางกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานของการวิจัย ทั้งนี้อาจเป็นเพราะ จำนวนข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และระหว่างกลุ่มข้อสอบที่มีระดับความง่ายปานกลางกับกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อใกล้เคียงกัน เมื่อนำจำนวนข้อสอบที่ลำเอียงมาทดสอบนัยสำคัญทางสถิติ จึงมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เมื่อตรวจสอบด้วยวิธีแมนเทิล-แฮนส์เซล พบว่า ข้อสอบที่มีความลำเอียงระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และ กลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานของการวิจัย ทั้งนี้อาจเป็นเพราะ จำนวนข้อสอบที่มีความลำเอียงที่ตรวจพบระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อไม่แตกต่างกันมากนัก เมื่อนำจำนวนข้อสอบที่ลำเอียงมาทดสอบนัยสำคัญทางสถิติ จึงมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เมื่อพิจารณาจำนวนข้อสอบที่มีความลำเอียงในแต่ละระดับความง่ายของข้อสอบ พบว่า กลุ่มข้อสอบที่มีระดับความง่ายต่ำ จะพบจำนวนข้อสอบที่มีความลำเอียงมากที่สุด คือ 4 ข้อ ซึ่งมีจำนวนใกล้เคียงกับกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง คือ จำนวน 3 ข้อ ส่วนกลุ่มข้อสอบที่มีระดับความง่ายสูงจะไม่พบจำนวนข้อสอบที่ลำเอียงเมื่อตรวจสอบด้วยวิธีชิปเทสท์ และตรวจพบได้เพียง 1 ข้อ โดยวิธีแมนเทิล-แฮนส์เซล แสดงว่า กลุ่มข้อสอบที่มีระดับความง่ายต่ำถึงปานกลาง ซึ่งเป็นข้อสอบที่ค่อนข้างยาก จะเกิดความลำเอียงต่อผู้สอบได้มากกว่ากลุ่มข้อสอบที่มีระดับความง่ายสูง ซึ่งเป็นข้อสอบที่ค่อนข้างง่าย และจะลำเอียงต่อกลุ่มผู้สอบเพศชายมากกว่ากลุ่มผู้สอบเพศหญิง

2. จำนวนข้อสอบที่มีความลำเอียง ระหว่างการตรวจสอบด้วยวิธีชิปเทสท์และวิธีแมนเทิล-แฮนส์เซล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีจำนวนข้อแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แสดงว่าพบจำนวนข้อสอบที่มีความลำเอียงใกล้เคียงกันระหว่างวิธีตรวจสอบทั้งสองวิธีในทุกระดับความ

ง่ายของข้อสอบ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานของการวิจัย ทั้งนี้อาจเป็นเพราะ ทั้งวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซลสามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ตรงกันทุกข้อในกลุ่มข้อสอบที่มีระดับความง่ายต่ำและกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และยังคงตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ใกล้เคียงกันในกลุ่มข้อสอบที่มีระดับความง่ายสูง เมื่อนำจำนวนข้อสอบที่ลำเอียงมาทดสอบนัยสำคัญทางสถิติ จึงมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้สอดคล้องกับงานวิจัยของ นารายานาน และสวามินาธาน (Narayanan and Swaminathan, 1994 : 315 – 328) ที่ศึกษาเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล โดยใช้ข้อมูลที่จำลองขึ้น พบว่า ทั้งวิธีชิปเทสต์ และวิธีแมนเทล-แฮนส์เซลสามารถตรวจพบข้อสอบที่มีความลำเอียงได้ใกล้เคียงกัน และยังคงสอดคล้องกับงานวิจัยของ พรณี จินตมาศ (2540 : 82) ที่ศึกษาเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีวิเคราะห์ 3 วิธี คือ วิธีแปลงค่าความยาก วิธีชิปเทสต์ และวิธีแมนเทล-แฮนส์เซล ในกลุ่มผู้สอบ 500 คน และ 1,000 คน ผลการวิจัยพบว่า จำนวนข้อสอบที่มีความลำเอียงจากการวิเคราะห์ความลำเอียงด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งในขนาดกลุ่มผู้สอบ 500 คน และ 1,000 คน

เมื่อพิจารณาจำนวนข้อสอบที่มีความลำเอียงระหว่างวิธีตรวจสอบความลำเอียงของข้อสอบแต่ละวิธี พบว่า ทั้งวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซลสามารถตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ตรงกันทุกข้อในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ และกลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และ ยังคงตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ใกล้เคียงกันในกลุ่มข้อสอบที่มีระดับความง่ายสูง แสดงว่า ทั้งวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล มีความไวในการตรวจพบจำนวนข้อสอบที่มีความลำเอียงได้ใกล้เคียงกัน

3. ค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง เมื่อตรวจสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานของการวิจัย ทั้งนี้อาจเป็นเพราะ ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้วจากการตรวจสอบทั้งวิธีชิปเทสต์และวิธีแมนเทลแฮนส์เซล ระหว่างกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูงมีค่าใกล้เคียงกัน เมื่อนำมาทดสอบนัยสำคัญทางสถิติ จึงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่สอดคล้องกับงานวิจัยของ ยศอารีย์ รวยชนพานิช (2534:56) ที่ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยากและความยาวของแบบทดสอบต่างกัน ผลการวิจัย

พบว่า แบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยากต่างกัน เมื่อมีจำนวนข้อเท่ากัน มีค่าความเชื่อมั่น ไม่แตกต่างกัน และแบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยาก 0.20 – 0.80 เมื่อมีจำนวนข้อต่างกัน มีค่าความเชื่อมั่น ไม่แตกต่างกัน

4. ค่าความเชื่อมั่นของแบบทดสอบ หลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เชล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำ กลุ่มข้อสอบที่มีระดับความง่ายปานกลาง และกลุ่มข้อสอบที่มีระดับความง่ายสูง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งผลที่ได้ไม่สอดคล้องกับสมมติฐานของการวิจัย ทั้งนี้อาจเป็นเพราะ ค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว ระหว่างการตรวจสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เชล ในกลุ่มข้อสอบที่มีระดับความง่ายต่ำและกลุ่มข้อสอบที่มีระดับความง่ายปานกลางมีค่าเท่ากัน ส่วนในกลุ่มข้อสอบที่มีระดับความง่ายสูงนั้นมีค่าใกล้เคียงกัน เมื่อนำมาทดสอบนัยสำคัญทางสถิติ จึงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่สอดคล้องกับงานวิจัยของ พรณี จินตมาศ (2540 : 86) ที่ศึกษาเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีวิเคราะห์ 3 วิธี คือ วิธีแปลงค่าความยาก วิธีชิปเทสต์ และวิธีแมนเทล-แฮนส์เชล ในกลุ่มผู้สอบ 500 คน และ 1,000 คน ซึ่งพบว่า ค่าความเชื่อมั่นของแบบทดสอบจากการวิเคราะห์ความลำเอียงด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เชล มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งในขนาดกลุ่มผู้สอบ 500 คน และ 1,000 คน

ข้อเสนอแนะในการวิจัย

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

จากการวิจัย พบว่า ข้อสอบที่ค่อนข้างยากจะเกิดความลำเอียงต่อผู้สอบมากกว่าข้อสอบที่ค่อนข้างง่าย และจะเกิดความลำเอียงต่อผู้สอบเพศชายมากกว่าผู้สอบเพศหญิง ดังนั้น ในการสร้างแบบทดสอบวัดความสามารถด้านภาษา จึงควรคัดเลือกเฉพาะข้อสอบที่ค่อนข้างง่ายหรือไม่ยากจนเกินไป เพื่อให้เกิดความยุติธรรมในการสอบ

2. ข้อเสนอแนะในการทำวิจัย

2.1 ควรมีการศึกษาในทำนองเดียวกันนี้ ในระดับชั้นหรือเนื้อหาวิชาที่แตกต่างกัน
ออกไป

2.2 ควรมีการศึกษาในทำนองเดียวกันนี้ โดยใช้การตรวจสอบความลำเอียงของข้อสอบด้วยวิธีอื่น ๆ และหาความสอดคล้องกับผลการศึกษาในครั้งนี้

บรรณานุกรม

บรรณานุกรม

- คมศักดิ์ ชื่นชม. (2539). การศึกษาผลการวิเคราะห์ความลำเอียงที่ใช้วิธีต่างกันของแบบทดสอบ
วัดจริยธรรมด้านความซื่อสัตย์. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ :
บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- จิตรสุดา ธารพร. (2539). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบจากแบบ
ทดสอบโจทย์ปัญหาทางคณิตศาสตร์และแบบทดสอบวัดความเข้าใจทางภาษา.
ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรี
นครินทรวิโรฒ. ถ่ายเอกสาร.
- เจน จันทรแสง. (2537). การวิเคราะห์ความลำเอียงของข้อสอบของแบบทดสอบวัดผลสัมฤทธิ์ทาง
การเรียนรู้คณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา).
กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร. ถ่ายเอกสาร.
- ซัชชัย เผ่าพงษ์. (2527). การวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบวัดความถนัดทาง
การเรียนรู้คณิตศาสตร์และภาษา ระดับชั้นมัธยมศึกษาตอนต้น. วิทยานิพนธ์ ศศ.ม.
(การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น. ถ่ายเอกสาร.
- ชวาล แพร์ตกุล และคนอื่น ๆ. (2514). หลักสูตรอบรมพิเศษวิชาวัดผลการศึกษาระยะสั้น.
กรุงเทพฯ : สำนักทดสอบและจิตวิทยาทางการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ชวาล แพร์ตกุล. (2518). เทคนิคการวัดผล. กรุงเทพฯ : วัฒนาพานิช.
- _____. (2530). เทคนิคการเขียนคำถามเลือกตอบ. กรุงเทพฯ : กิ่งจันทร์การพิมพ์.
- ญาณภัทร สีหะมงคล. (2540). การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำ
หน้าที่ต่างกันระหว่าง วิธี Lord's χ^2 วิธี Raju's Area Measures และวิธี Closed Interval
Area. ปรินูญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ : บัณฑิต
วิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ทัศนีย์ พิรมนตรี. (2530). การวิเคราะห์ความลำเอียงของแบบทดสอบวิชาคณิตศาสตร์
โครงการตรวจสอบคุณภาพการศึกษา ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526.
วิทยานิพนธ์ ค.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- นิรมล ชัยชลิต. (2537). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของแบบทดสอบวัด
ความเข้าใจในการอ่านภาษาไทย ตามทฤษฎีคลาสสิกอลที่ใช้วิธีวิเคราะห์ต่างกัน.
ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย
ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- นวลอนงค์ นันทวงศ์. (2543). การเปรียบเทียบความลำเอียงของข้อสอบของแบบทดสอบความเข้าใจในการอ่าน ตามลักษณะการใช้ภาษาพูดในครอบครัว โดยวิธีวิเคราะห์ต่างกัน. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พัชรี ปิยภักดิ์. (2531). การวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ชั้นประถมศึกษาปีที่ 6. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พรรณี จินตมาศ. (2540). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้ขนาดกลุ่มผู้สอบและวิธีวิเคราะห์ต่างกัน. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ยศอารีย์ รวยชนพานิช. (2534). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยากและความยาวของแบบทดสอบต่างกัน ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- เรวดี อินทสระ. (2539). ผลการตรวจสอบความลำเอียงของข้อสอบต่อการศึกษาความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน. ปรินูญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ล้วน สายยศ และ อังคณา สายยศ. (2536). เทคนิคการวิจัยทางการศึกษา. พิมพ์ครั้งที่ 3. กรุงเทพฯ : ศูนย์ส่งเสริมวิชาการ.
- _____ . (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ: ชรรมเด็ก.
- อนันต์ ศรีโสภ. (2525). ทฤษฎีการวัดและการทดสอบ. กรุงเทพฯ: ไทยวัฒนาพานิช.
- อังคณา สายยศ. (2539). การศึกษาเปรียบเทียบความลำเอียงของแบบทดสอบวัดความเข้าใจในการอ่านตามลักษณะการใช้ภาษาพูดในครอบครัว. กรุงเทพฯ : ภาควิชาการวัดผลและวิจัยทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- อารี วัชรโสถธิกุล. (2543). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้รูปแบบและวิธีการต่างกัน. ปรินูญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- Allen, M.J. and Yen, W.M. (1979). *Introduction to Measurement Theory*. Monterey : California Brooks / Cole.
- Berk, R. A. (Ed.) (1982). *Handbook of Method for Detecting Test Bias*. Bartimore and London MD : The Johns Hopkin University Press.
- Burrill, L.E. (1982). "Comparative Studies of Item Bias Method," In *Handbook of Methods for Detecting Test Bias*. Baltimore and London : The Johns Hopkins University Press.
- Camilli, Gregory. and Shepard, Lorrie A. (1994). *Methods for Identifying Biased Test Items*. California : Sage Publications.
- Ferguson, G.A. (1981). *Statistical Analysis in Psychology and Education*. New York:Mc Graw - Hill.
- Freedle, Roy. and Kostin, Irene. (1990,Winter). "Item Difficulty of Four Verbal Item Types and an Index of Differential Item Functioning for Black and White Examinees," *Journal of Educational Measurement*. 27(4) : 329 – 343.
- Hambleton, Ranauld K., H. and Swaminathan and Roger H. Jame. (1991). *Foudamemals of item Response Theory*. Newbury Park, california : Sage Publications.
- Intasuwan, P. (1979). *A Comparison of Three Approaches for Determining item Bias in Cross National Testing*. Dissertation Ph.D. Pittsburgh : Graduate School, The University of Pittsburgh. PhotoCoppied.
- Jensen, Arthur. R. (1980). *Bias in Metal Testing*. Newyork : A division of Macmilla Publising.
- Lord, F.M. and Novick , M.R. (1968). *Statistical Theories of Mental Test Scores*. Reading MA : Addison-Wesley. Publishing Company.
- Mazor, K.M., Clauser, B.E. and Hambleton, R.K. (1992,Summer). "The Effect of Sample Size on the functioning of the Mantel-Haenzel staticce," *Educational and Psychological Measurement*. 52(2) : 443 - 451.
- Mc Canley, D. Cynthia and Joge Mendoza. (1985,December). "A Simulation Study Item Bias Using a Two-Parameter Item Response Model," *Applied Psychological Measurement*. 9(4) : 389 - 400.
- Miller, M.D. and Oshima, T.C. (1992,December). "Effect of Sample Size, number of Biased Items, and magnitude of Bias on a Two-Stage Item Bias estimation Method," *Applied Psychological measurement*. 16(4) : 381 – 388.

- Millsap, R. E. and Everson, H.T. (1993, December). "Methodology Review : Statistical Approaches for Assessing Measurement Bias," *Applied Psychological Measurement*. 17(4) : 297 - 334.
- Narayanan, Pankaja and Swaminathan, H. (1994, December). "Performance of the Mantel-Haenszel and Simultaneous Item Bias Procedures for Detecting Differential Item function," *Applied Psychological Measurement*. 18(4) : 315 – 328.
- Roussos, Louis A. and Stout, William F. (1996, Summer). "Simulation Studies of the Effect of Small Sample Size and Studies Item Parameters on SIBTEST and Mantel-Haenszel Type I Error Performance," *Journal of Educational Measurement*. 33(2) : 215 – 230.
- Shealy, Robin and Stout William. (1993, June). "A Model-Based Standardization Approach that Separates True Bias/DIF from group Ability Difference and Detects Test Bias/DIF as Well as Item Bias/DIF," *Psychometrika*. 58(2) : 159 – 194.
- Snedecor and Cochran. (1967). *Statistical Method*. Iowa : Iowa State University Press.
- Subkoviak, Micheal J., et al. (1984, Spring). "Empirical Comparison of Detected Item Bias Detection Procedure with Bias Manipulation," *Journal of Educational Measurement*. 21(10) : 49 - 59.

ภาคผนวก

ภาคผนวก ก

- เครื่องมือที่ใช้ในการศึกษาค้นคว้า
- สรุปจำนวนข้อของแบบทดสอบวัดความสามารถด้านภาษา
จำแนกตามระดับความง่ายของข้อสอบ

แบบทดสอบวัดความสามารถด้านภาษา

คำชี้แจง

1. แบบทดสอบฉบับนี้เป็นแบบทดสอบวัดความสามารถด้านภาษาโดยเป็นแบบทดสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก มีจำนวน 60 ข้อ ใช้เวลาในการทำ 40 นาที โดยแบ่งเป็น 5 ด้าน ได้แก่

- ด้านคำตรงข้าม จำนวน 10 ข้อ
- ด้านคำที่มีความหมายใกล้เคียง จำนวน 10 ข้อ
- ด้านคำศัพท์สัมพันธ์ จำนวน 10 ข้อ
- ด้านความหมายไม่เข้าพวก จำนวน 10 ข้อ
- ด้านความเข้าใจภาษา จำนวน 20 ข้อ

2. วิธีการตอบ ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุดหรือเหมาะสมที่สุดเพียงข้อละ

1 คำตอบ โดยทำเครื่องหมายกากบาท (X) ลงในกระดาษคำตอบที่แจกให้

3. ห้ามนักเรียนขีดเขียนหรือทำเครื่องหมายใด ๆ ลงในข้อสอบ

วิธีการตอบ

- เมื่อต้องการตอบข้อ ก ให้ทำดังนี้

ข้อ	ก	ข	ค	ง	จ
00	X				

- เมื่อต้องการเปลี่ยนคำตอบเป็น จ ให้ทำดังนี้

ข้อ	ก	ข	ค	ง	จ
00	X				X

ขอขอบคุณในความร่วมมือ

ด้านคำตรงข้าม

คำชี้แจง ให้นักเรียนพิจารณาหาคำที่มีความหมายตรงข้ามกับคำที่ขีดเส้นใต้ในแต่ละข้อ

1. ร่ำรวย

- ก. อี๊ดอึด
- ข. เย็นชา
- ค. วางเฉย
- ง. ซึมเศร้า
- จ. เจียบขรึม

2. คล่องแคล่ว

- ก. นุ่มน้ม
- ข. เชื่องช้า
- ค. อี๊ดอาด
- ง. เนิบนาบ
- จ. เฉื่อยชา

3. เข้มนั่น

- ก. จืดชืด
- ข. ละลาย
- ค. บางเบา
- ง. เจือจาง
- จ. เลือนราง

4. พยาบาท

- ก. ให้อภัย
- ข. ขอโทษ
- ค. ชื่นชม
- ง. ยอมรับ
- จ. ปลอ่ยวาง

5. เขาพูดจาไพเราะ

- ก. โกหก
- ข. ไร้สาระ
- ค. พุดมาก
- ง. เหลวไหล
- จ. หยาบคาย

6. เขาเป็นคนหูเบา

- ก. นึกคิด
- ข. มั่นใจ
- ค. เข้มแข็ง
- ง. หนักแน่น
- จ. ระมัดระวัง

7. นักเรียนห้องนี้มีความสามัคคีกัน

- ก. พินาศ
- ข. ทะเลาะ
- ค. ห่างเหิน
- ง. แดกแยก
- จ. บาดหมาง

8. ครอบครัวของเขาสมบูรณ์พูนสุขดี

- ก. ลุ่ม ๆ ดอน ๆ
- ข. ระหกระเหิน
- ค. พอยู่พอกิน
- ง. ขัดสนจนยาก
- จ. เดือดร้อนวุ่นวาย

9. เขาเป็นคนมีระเบียบในการทำงาน

- ก. ละเลย
- ข. ขี้เกียจ
- ค. เพิกเฉย
- ง. มั่งง่าย
- จ. หย่อนยาน

10. คุณอย่าเพิ่งเร่งรัดฉันได้ไหม

- ก. จู้จู้
- ข. ชะลอ
- ค. ระงับ
- ง. หยุดชะงัก
- จ. หน่วงเหนี่ยว

ด้านคำที่มีความหมายใกล้เคียง

คำชี้แจง ให้นักเรียนพิจารณาหาคำที่มีความหมายใกล้เคียงมากที่สุดกับคำที่ขีดเส้นใต้ในแต่ละข้อ

11. แก้ไข

- ก. ตกแต่ง
- ข. ปรับปรุง
- ค. เพิ่มเติม
- ง. ซ่อมแซม
- จ. เปลี่ยนแปลง

12. อึกทึก

- ก. จอแจ
- ข. สับสน
- ค. คับคั่ง
- ง. คึกคัก
- จ. เสียงดัง

13. มโหฬาร

- ก. โอ่อ่า
- ข. โชติช่วง
- ค. ยิ่งใหญ่
- ง. มากมาย
- จ. กว้างขวาง

14. อัปลักษณ์

- ก. พิสดาร
- ข. น่ากลัว
- ค. น่าเกลียด
- ง. เหลวไหล
- จ. ประหลาด

15. การทำงานบางครั้งต้องมีอุปสรรค

- ก. ศัตรู
- ข. ปัญหา
- ค. เลื่อนหนาม
- ง. ความทุกข์
- จ. สิ่งกีดขวาง

16. เขาเป็นคนมีความพยายาม

- ก. ขยัน
- ข. ใฝ่รู้
- ค. มุ่งมั่น
- ง. อุตสาหะ
- จ. ทะเยอทะยาน

17. หยุดนะ! อย่าขยับ

- ก. ยุกยิก
- ข. ขยาย
- ค. หายใจ
- ง. เคลื่อนไหว
- จ. เปลี่ยนแปลง

18. เขามีที่ทำนั้งกับฉัน

- ก. ถือตัว
- ข. เย็นชา
- ค. แข็งอน
- ง. หงุดหงิด
- จ. หลงตัวเอง

19. จะมานั่งคร่ำครวญอยู่ทำไม

- ก. รำพัน
- ข. ร้องไห้
- ค. ซึมเศร้า
- ง. ลำเอียง
- จ. หม่นหมอง

20. ผู้หญิงคนนี้มีกริยามารยาทนุ่มนวล

- ก. สุภาพ
- ข. งดงาม
- ค. ชดช้อย
- ง. อ้อยส้อย
- จ. อ่อนหวาน

ด้านคำศัพท์สัมพันธ์

คำชี้แจง ให้นักเรียนพิจารณาหาคำที่สัมพันธ์หรือเกี่ยวข้องมากที่สุดกับคำที่กำหนดให้ในแต่ละข้อ

- | | |
|---|---|
| <p>21. ไสยศาสตร์</p> <p>ก. ผีसाง</p> <p>ข. ความกลัว</p> <p>ค. ความลึกลับ</p> <p>ง. การพิสูจน์</p> <p>จ. เครื่องราง</p> <p>22. ออกกำลังกาย</p> <p>ก. กีฬา</p> <p>ข. เหยื่อ</p> <p>ค. สนุก</p> <p>ง. สามัคคี</p> <p>จ. สุขภาพ</p> <p>23. ลูกสักหลาด</p> <p>ก. กอล์ฟ</p> <p>ข. เทนนิส</p> <p>ค. ปิงปอง</p> <p>ง. แบดมินตัน</p> <p>จ. บาสเก็ตบอล</p> <p>24. รั้วของชาติ</p> <p>ก. ครู</p> <p>ข. หมอ</p> <p>ค. ตำรวจ</p> <p>ง. ทหาร</p> <p>จ. ข้าราชการ</p> <p>25. กลางคืน</p> <p>ก. ท้องฟ้า</p> <p>ข. นอนหลับ</p> <p>ค. ความฝัน</p> <p>ง. ความมืด</p> <p>จ. ความเงียบ</p> | <p>26. หนังสือพิมพ์</p> <p>ก. ข่าว</p> <p>ข. กีฬา</p> <p>ค.บันเทิง</p> <p>ง. สารคดี</p> <p>จ. มาตรการ</p> <p>27. บุหรี่</p> <p>ก. ใต้</p> <p>ข. ตับ</p> <p>ค. ปอด</p> <p>ง. หัวใจ</p> <p>จ. กระเพาะ</p> <p>28. เจ้าพ่อ</p> <p>ก. อาวุธ</p> <p>ข. ตำรวจ</p> <p>ค. อิทธิพล</p> <p>ง. ลูกน้อง</p> <p>จ. กฎหมาย</p> <p>29. อุณหภูมิ</p> <p>ก. อากาศ</p> <p>ข. แสงแดด</p> <p>ค. แสงสว่าง</p> <p>ง. ความร้อน</p> <p>จ. ความเร็ว</p> <p>30. มนุษย์สัมพันธ์</p> <p>ก. ตรงต่อเวลา</p> <p>ข. เข้ากับผู้อื่นได้ดี</p> <p>ค. อ่อนน้อมถ่อมตน</p> <p>ง. ไม่เอาเปรียบผู้อื่น</p> <p>จ. มีความรับผิดชอบ</p> |
|---|---|

ด้านความหมายไม่เข้าพวก

คำชี้แจง ให้นักเรียนพิจารณาว่า ในแต่ละข้อต่อไปนี้ ข้อใดมีความหมายแตกต่างจากข้ออื่น ๆ

31. ก. ทำจิตใจให้สบาย
ข. ออกกำลังกายทุกวัน
ค. แต่งกายให้ถูกต้องเหมาะสม
ง. รับประทานอาหารครบ 5 หมู่
จ. พักผ่อนนอนหลับให้เพียงพอ

32. ก. ลื่นกับพื้น
ข. ขมื่นกับปูน
ค. เกลือจิ้มเกลือ
ง. จิงกีรา ข่าก็แรง
จ. น้ำพิ่งเรือ เสือพิ่งป่า

33. ก. สรงน้ำพระ
ข. รดน้ำดำหัว
ค. แห่เทียนพรรษา
ง. สาคน้ำสรงกรานต์
จ. ปลอยนกปล่อยปลา

34. ก. เขาลืมตา
ข. อย่าลืมฉัน
ค. ฉันลืมคืนหนังสือ
ง. ลืมส่งจดหมาย
จ. อย่าลืมทานยา

35. ก. ป้อมกำลังดูบ้าน
ข. สุดากำลังซักผ้า
ค. แม่กำลังทำกับข้าว
ง. แดงกำลังรดน้ำต้นไม้
จ. พ่อกำลังอ่านหนังสือพิมพ์

36. ก. ไปลามาไหว้
ข. มีสัมมาคารวะ
ค. รู้จักกาลเทศะ
ง. รักสวยรักงาม
จ. อ่อนน้อมถ่อมตน

37. ก. คนลืมตัว
ข. วัวลืมดิน
ค. กิ้งก่าได้ทอง
ง. คางคกขึ้นวอ
จ. หัวล้านได้หวี

38. ก. คิดดีแล้วหรือ
ข. อย่าควั่นคัตสันใจ
ค. ตกลงตามนี้แหละ
ง. มันเสียงเกินไป
จ. เดียวก่อนก็ได้

39. ก. ปอไม่เคยพูดโกหก
ข. ปลาไม่ชอบทำร้ายสัตว์
ค. ปูไม่เคยขโมยของผู้อื่น
ง. ปีนไม่ชอบดื่มของมึนเมา
จ. ป่านไม่เคยลอกการบ้านเพื่อน

40. ก. นพเป็นคนตรงไปตรงมา
ข. นัทเป็นคนชอบพูดเปิดเผย
ค. หน้อยชอบพูดหวานผ่าซาก
ง. แนนอยากพูดอะไรก็พูดออกมา
จ. นื่องไม่เคยพูดทำร้ายจิตใจใคร

ด้านความเข้าใจภาษา

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 41 - 43

“ผมยอมรับว่า จากการเลี้ยงหนูตะเภาเป็นเวลา 8 เดือน กับการเลี้ยงหมา 2 อาทิตย์ หมาให้ความรู้สึกที่ลึกยิ่งกว่าหลายเท่า “หมา” เราไม่ต้องทำใจให้รักมัน พฤติกรรมที่แสดงออกของลูกหมา การพุ่งเข้าใส่ ไล่จับมือ ทำทางการเดิน ทำให้เราเกิดความรักและผูกพันอย่างง่ายดาย ผิดกับหนูตะเภาที่เราต้องทำใจให้รักมันเอง มันทำได้แค่กิน และก้มหน้าก้มตากินอย่างเดียว”

- | | |
|---|---|
| <p>41. ผู้เขียนคิดว่า หนูตะเภากับสุนัข มีความแตกต่างกันที่สิ่งใด</p> <p>ก. การกินอาหาร</p> <p>ข. การเคลื่อนไหว</p> <p>ค. ความเฉลียวฉลาด</p> <p>ง. ความน่ารักและขี้เล่น</p> <p>จ. พฤติกรรมที่แสดงออก</p> | <p>42. จากข้อความที่ว่า <u>หมาให้ความรู้สึกที่ลึกยิ่งกว่าหลายเท่า</u> หมายถึงความรู้สึกใด</p> <p>ก. เอ็นดู</p> <p>ข. ห่วงใย</p> <p>ค. ผูกพัน</p> <p>ง. ห่วงเหิน</p> <p>จ. อยากเป็นเจ้าของ</p> |
|---|---|

43. จากข้อความนี้ อาจสรุปได้ว่าอย่างไร

- ก. ชายคนนี้จะเลิกเลี้ยงหนูตะเภา
- ข. ชายคนนี้จะเลี้ยงแต่หมาอย่างเดียว
- ค. ชายคนนี้จะพยายามเลี้ยงหนูตะเภาต่อไป
- ง. ชายคนนี้จะรู้สึกรักหมามากกว่าหนูตะเภา
- จ. ชายคนนี้จะเลี้ยงสัตว์ทุกชนิดยกเว้นหนูตะเภา

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 44 – 45

“ภาพเจนตาในโรงเรียนบางครั้งก็สอนอะไรกับเราได้หลายอย่าง เช่น ภาพที่เด็กชายหญิงตัวน้อยวิ่งเซหลุน ๆ หกล้มหัวเข้ากระแทกพื้น ตอนแรกไม่ได้ยินทั้งเสียงร้องและไม่ได้เห็นน้ำตาแต่พอหันมองแม่เห็นตลกอกตกใจร้องกรี๊ดกร๊าดผลุนผลันจะเข้ามาช่วยเท่านั้นละเลยแขนอ่อนโยนอยากให้แม่เข้ามาประคองอย่างเคย แถมตะเบ็งเสียงร้องให้จำคุณแมียิ่งตกอกตกใจใหญ่ไปกว่าลูกเจ็บมากยิ่งขึ้น เห็นลูกเลือดออกซิบ ๆ เรื่องก็ดูจะยิ่งโกลาหลหนักข้อเข้าไปอีก”

- | | |
|--|--|
| <p>44. พฤติกรรมของเด็กจากข้อความนี้ ตรงกับข้อใด</p> <p>ก. ไร้เดียงสา</p> <p>ข. ซุกซนดื้อรั้น</p> <p>ค. ขาดความอบอุ่น</p> <p>ง. เอาแต่ใจตนเอง</p> <p>จ. เรียกร้องความสนใจ</p> | <p>45. เพราะเหตุใดเด็กจึงร้องไห้เมื่อเห็นแม่</p> <p>ก. ตกใจกลัว</p> <p>ข. เพิ่งรู้สึกเจ็บ</p> <p>ค. กลัวโดนแม่ตี</p> <p>ง. เห็นเลือดตนเอง</p> <p>จ. อยากให้แม่มาช่วย</p> |
|--|--|

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 46 – 48

“เกี่ยวกับในเรื่องนี้ ไม่ใช่ปัญหายุ่งยากแต่อย่างใด เรามีเคล็ดลับมาฝากกันอีกเช่นเคย ด้วยการนำบะหมี่ที่เหลือมาห่อด้วยใบผักกาดเขียวหรือผักกาดกวางตุ้ง แต่ต้องเน้นที่สด ๆ เหี่ยว ๆ ไม่เอาเมื่อห่อเสร็จแล้วนำไปห่อด้วยผ้าเป็นชั้นที่สอง และห่อด้วยผักกาดเขียวหรือผักกาดกวางตุ้งเป็นชั้นที่สามอีกครั้ง ต่อจากนั้นจึงนำเข้าแช่ตู้เย็นในช่องแช่ผัก เพียงเท่านั้นบะหมี่ของท่านก็จะอยู่ยงคงกระพันสามารถเก็บไว้ได้นานนับเดือนเชียวนะ”

46. คำว่า คงกระพัน หมายถึงคุณสมบัติใด

- ของบะหมี่
- ก. ชุ่มชื้น
- ข. เหนียวนุ่ม
- ค. เก็บได้นาน
- ง. สดใหม่อยู่เสมอ
- จ. รสชาติกลมกล่อม

47. เพราะเหตุใด จึงนำใบผักกาดเขียวหรือ

- ผักกาดกวางตุ้งมาห่อบะหมี่
- ก. เหี่ยวช้า
- ข. ใบมีสีเขียว
- ค. หาซื้อได้ง่าย
- ง. ใบมีขนาดใหญ่
- จ. รักษาความชุ่มชื้นได้ดี

48. ควรตั้งชื่อข้อความนี้ว่าอย่างไรจึงจะเหมาะสม

- ก. วิธีการถนอมเส้นบะหมี่
- ข. การปรุงอาหารด้วยบะหมี่
- ค. เก็บบะหมี่อย่างไรให้ถูกวิธี
- ง. การห่อบะหมี่ด้วยผักใบเขียว
- จ. ลวกบะหมี่อย่างไรให้เส้นเหนียวนุ่ม

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 49 – 50

“ไฟแดงสลับเขียวและเขียวสลับแดงอยู่เช่นนั้นเป็นซ้ำซาก คำจึงเริ่มมีโอกาสเช็ดรถโดยเจ้าของไม่คิดตะเพิดอยู่หลายคัน แคมบางคันยังบ่นคำวน่าสงสารให้ได้ยิน คำได้สวดคำมาคันละห้าบาทบ้างสิบบาทบ้าง บางคันเช็ดจนเพลินพอเจ้าของรถเตรียมเปิดกระจกจะให้สวดคำก็กลับเปลี่ยนใจปั้งรถออกไปเพราะไฟเขียว”

49. การทำงานของคำต้องอาศัยสิ่งใดมากที่สุด

- ก. ความขยัน
- ข. ความอดทน
- ค. ความซื่อสัตย์
- ง. ความคล่องแคล่ว
- จ. ความระมัดระวัง

50. ข้อใดสรุปไม่ถูกต้องเกี่ยวกับอาชีพของคำ

- ก. ต้องทำงานตลอดทั้งวัน
- ข. ในแต่ละวันได้เงินไม่เท่ากัน
- ค. ต้องผจญกับฝุ่นและควันพิษ
- ง. ต้องเสี่ยงอันตรายจากอุบัติเหตุ
- จ. คำทำงานอยู่ตามสี่แยกไฟแดง

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 51 – 54

“จากนี้ไปอีกพันล้านปี ดวงอาทิตย์จะยิ่งเปล่งแสงแรงกล้ายิ่งขึ้นกว่าเดิมอีก 10% โดยการจำลองแบบขึ้นดูแล้วว่า โลกของเราที่แรกจะเปลี่ยนไปมีสภาพเหมือนกับเรือนกระจกสำหรับปลูกต้นไม้อันระอุอ้าวและเมื่อระยะเวลาผ่านไปอีกสามพันล้านปี ดวงอาทิตย์จะยิ่งแผดแสงแรงกล้านักขึ้นเป็น 40% ซึ่งหมายถึง อดสูของชีวิตบนดาวเคราะห์ดวงนี้ เพื่อที่จะหลบเลี่ยงจากความร้อนแรงกล้านี้อยู่ทางเดียวเท่านั้น คือ จะต้องให้โลกเคลื่อนที่ด้วยอัตราเร็วของวงโคจรเสียใหม่ให้ยาวไกลออกไป”

51. คำว่า ดาวเคราะห์ดวงนี้ หมายถึงข้อใด

- ก. โลก
- ข. ดาวศุกร์
- ค. ดวงจันทร์
- ง. ดวงอาทิตย์
- จ. ดาวอังคาร

52. จากข้อความนี้ คาดว่าจะเกิดเหตุการณ์ใดในอนาคต

- ก. โลกจะแตกสลาย
- ข. สิ่งมีชีวิตจะสูญพันธุ์
- ค. ดวงอาทิตย์จะระเบิดปะทุ
- ง. ดวงอาทิตย์จะมีขนาดใหญ่ขึ้น
- จ. โลกจะเคลื่อนที่ออกห่างดวงอาทิตย์

53. วิธีแก้ไขปัญหามาจากข้อความนี้ คือข้อใด

- ก. ช่วยกันปลูกต้นไม้ให้มาก ๆ
- ข. ช่วยกันรักษาสภาพแวดล้อม
- ค. มนุษย์ต้องย้ายไปอยู่บนดวงจันทร์
- ง. ผลักดันดวงอาทิตย์ให้ออกห่างโลก
- จ. เคลื่อนย้ายโลกไปอยู่ในวงโคจรใหม่

54. จากข้อความนี้ น่าจะเป็นข้อมูลที่ได้มาจากบุคคลกลุ่มใด

- ก. นักดาราศาสตร์
- ข. นักโบราณคดี
- ค. นักสถิติศาสตร์
- ง. นักวิทยาศาสตร์
- จ. นักทรัพยากรธรณี

คำชี้แจง จากข้อความต่อไปนี้ จงตอบคำถามข้อ 52 – 56

“ผมยังมีสติตั้งค์เต็มร้อยครับ แต่ที่ต้องอ่านโฆษณาอย่างเป็นบ้าเป็นหลังเช่นนั้น เพราะผมสุดจะทนกับประสบการณ์ที่ซ้ำซากเกี่ยวกับบ้านเช่าที่นำแห่งหน้าและกำลังจะหาบ้านสักหลัง บ้านที่จะหานั้นเป็นบ้านที่ซื้อโดยมีผมเป็นเจ้าของ แต่ต้องเป็นบ้านที่มีขนาดกระป๋องที่ราคาสามัคคีกับเงินในกระเป๋าของผมด้วย”

55. จากข้อความ อ่านโฆษณาอย่างเป็นบ้าเป็นหลัง คำที่ขีดเส้นใต้ หมายถึงข้อใด

- ก. ไม่รู้สึกตัว
- ข. เอาจริงเอาจัง
- ค. พินิจพิจารณา
- ง. อ่านซ้ำแล้วซ้ำอีก
- จ. ควบคุมอารมณ์ไม่ได้

56. จากข้อความ ราคาสามัคคีกับเงินในกระเป๋า คำว่า สามัคคี หมายถึงข้อใด

- ก. คู่มีค่า
- ข. ถูกกว่า
- ค. ประหยัด
- ง. ใกล้เคียง
- จ. เท่าเทียมกัน

คำชี้แจง อ่านบทประพันธ์ต่อไปนี้ แล้วตอบคำถามข้อ 57 – 60

“ทั้งนี้เป็นต้นด้วยผลเหตุ
มิได้พิจารณาข้าไท
ไม่รู้รอบประกอบในราชกิจ
สุภายิตท่านกล่าวเป็นรามา
ไม่ควรอย่าให้อัครฐาน
เพราะไม่เชื่อด่านานโบราณคดี

จะอาเพศเพราะผู้เป็นใหญ่
เคยใช้ก็เลี้ยงด้วยเมตตา
ประพฤติกการแต่ผิดด้วยอจนา
จะแต่งตั้งเสนาบดี
จะเสียการแผ่นดินกรุงศรี
จึงเสียทิวาศักยตรา”

57. คำประพันธ์ข้างต้น กล่าวถึงเรื่องใด

- ก. การรักษากรุงราช
- ข. การทำศึกสงคราม
- ค. ความเมตตาการุณา
- ง. การแต่งตั้งข้าราชการ
- จ. การปกครองบ้านเมือง

58. เสนาบดี เปรียบได้กับตำแหน่งใดใน

- ปัจจุบัน
- ก. นักการเมือง
- ข. นายกรัฐมนตรี
- ค. สมาชิกวุฒิสภา
- ง. สมาชิกสภาผู้แทนราษฎร
- จ. รัฐมนตรีว่าการกระทรวง

59. เพราะเหตุใด จึงแต่งตั้งข้าทาสบริวารเป็น
เสนาบดี

- ก. หาบุคคลอื่นที่เหมาะสมไม่ได้
- ข. ประหยัดงบประมาณรายจ่าย
- ค. เพื่อให้ข้าทาสบริวารได้ดิบได้ดี
- ง. เพื่อความสะดวกในการสั่งงาน
- จ. ไม่ต้องการให้คนภายนอกเข้ามายุ่ง
เกี่ยว

60. บทประพันธ์นี้ ให้ข้อคิดตรงกับข้อใด

- ก. ค่าของคนอยู่ที่ผลของงาน
- ข. ควรยึดถือคำสั่งสอนในอดีต
- ค. คิดจะทำงานใหญ่ต้องใจเย็น
- ง. พิจารณามูลค่าให้เหมาะสมกับงาน
- จ. อย่าไว้ใจทาง อย่าวางใจคน จะสนใจ
เอง

.....

ตาราง 9 สรุปจำนวนข้อของแบบทดสอบวัดความสามารถด้านภาษา จำแนกตามระดับความง่ายของข้อสอบ

แบบทดสอบวัดความสามารถด้านภาษา	ระดับความง่ายของข้อสอบ			รวม
	ต่ำ	ปานกลาง	สูง	
ด้านคำตรงข้าม	3	3	4	10
ด้านคำที่มีความหมายใกล้เคียง	4	3	3	10
ด้านคำศัพท์สัมพันธ์	3	4	3	10
ด้านความหมายไม่เข้าพวก	2	3	5	10
ด้านความเข้าใจภาษา	8	7	5	20
รวม	20	20	20	60

ภาคผนวก ข

ค่าความง่าย(p_E) และค่าอำนาจจำแนก(r) ของแบบทดสอบ

ตาราง 10 ค่าความง่าย (p_E) และ ค่าอำนาจจำแนก (r) ของแบบทดสอบวัดความสามารถด้านภาษา
จำนวน 60 ข้อ

ข้อ	ตัวเลือก	p_E	r	Δ	ข้อ	ตัวเลือก	p_E	r	Δ
1	ก	.15	.67	17.1	6	ก	.10	.02	18.0
	ข	.05	.39	19.7		ข	.18	.12	16.7
	ค	.05	.39	19.7		ค	.21	.30	16.3
	ง*	.69	.58	11.0		ง*	.39	.42	14.1
	จ	.10	.21	18.1		จ	.15	.00	17.1
2	ก	.05	.39	19.7	7	ก	.11	.49	17.9
	ข*	.62	.33	11.8		ข	.23	.19	16.0
	ค	.13	.08	17.5		ค	.05	.39	19.7
	ง	.05	.39	19.7		ง*	.43	.61	13.7
	จ	.19	.30	16.5		จ	.17	.40	16.8
3	ก	.26	.10	15.6	8	ก	.14	.37	17.4
	ข	.14	.04	17.3		ข	.16	.09	17.1
	ค	.07	.00	18.9		ค	.15	.21	17.2
	ง*	.36	.23	14.4		ง*	.48	.42	13.2
	จ	.16	.18	18.9		จ	.06	.16	19.1
4	ก*	.61	.45	11.8	9	ก	.11	.04	17.0
	ข	.06	.46	19.2		ข	.07	.09	18.8
	ค	.08	.05	18.6		ค	.32	.26	14.9
	ง	.08	.52	18.7		ง*	.47	.32	13.3
	จ	.14	.16	17.3		จ	.05	.39	19.7
5	ก	.09	.56	18.3	10	ก	.05	.39	19.7
	ข	.12	.62	17.7		ข	.14	.15	17.3
	ค	.06	.07	19.3		ค	.48	.07	13.2
	ง	.05	.39	19.7		ง	.14	.05	17.2
	จ*	.66	.65	11.4		จ*	.21	.25	16.2

* เป็นตัวเลือกที่ถูกต้อง

ตาราง 10 (ต่อ)

ข้อ	ตัวเลือก	p_E	r	Δ	ข้อ	ตัวเลือก	p_E	r	Δ
11	ก	.08	.17	18.7	16	ก	.40	.20	14.0
	ข*	.55	.24	12.5		ข	.13	.06	17.4
	ค	.14	.16	17.3		ค	.08	.26	18.5
	ง	.12	.12	17.7		ง*	.36	.35	14.5
	จ	.11	.04	17.0		จ	.05	.39	19.7
12	ก	.05	.39	19.7	17	ก	.13	.54	17.5
	ข	.12	.34	17.6		ข	.17	.23	16.9
	ค	.05	.39	19.7		ค	.05	.39	19.7
	ง	.22	.18	16.6		ง*	.38	.59	14.2
	จ*	.61	.37	11.9		จ	.26	.18	15.6
13	ก	.01	.24	18.2	18	ก	.06	.32	19.3
	ข	.07	.37	18.9		ข*	.35	.37	14.5
	ค*	.48	.54	13.2		ค	.18	.27	16.7
	ง	.16	.22	17.0		ง	.05	.39	19.7
	จ	.17	.31	16.9		จ	.38	.10	14.2
14	ก	.05	.39	19.7	19	ก*	.61	.32	11.8
	ข	.22	.32	16.2		ข	.05	.39	19.7
	ค*	.36	.37	14.4		ค	.12	.10	17.6
	ง	.05	.39	19.7		ง	.05	.39	19.7
	จ	.35	.03	14.5		จ	.20	.23	16.3
15	ก	.08	.17	18.7	20	ก	.11	.09	17.9
	ข*	.44	.35	13.6		ข	.10	.00	18.1
	ค	.05	.39	19.7		ค	.06	.47	19.1
	ง	.19	.29	16.6		ง	.05	.39	19.7
	จ	.23	.00	16.0		จ*	.71	.24	10.7

* เป็นตัวเลือกที่ถูกต้อง

ตาราง 10 (ต่อ)

ข้อ	ตัวเลือก	p _E	r	Δ	ข้อ	ตัวเลือก	p _E	r	Δ
21	ก	.09	.10	18.4	26	ก*	.62	.80	11.8
	ข	.14	.09	17.2		ข	.20	.73	16.4
	ค*	.39	.26	14.1		ค	.08	.53	18.6
	ง	.08	.19	18.5		ง	.07	.26	19.1
	จ	.28	.11	15.3		จ	.05	.39	19.7
22	ก	.05	.16	19.6	27	ก	.05	.39	19.7
	ข	.05	.39	19.7		ข	.07	.39	18.8
	ค	.21	.10	16.2		ค*	.56	.61	12.4
	ง	.41	.13	13.9		ง	.19	.41	16.5
	จ*	.34	.22	14.7		จ	.14	.37	17.4
23	ก	.05	.39	19.7	28	ก	.05	.39	19.7
	ข*	.43	.28	13.7		ข	.10	.57	18.1
	ค	.05	.39	19.7		ค*	.61	.65	11.9
	ง	.16	.14	17.0		ง	.19	.48	16.6
	จ	.39	.20	14.1		จ	.06	.24	19.2
24	ก	.05	.39	19.7	29	ก	.14	.16	17.3
	ข	.05	.39	19.7		ข	.15	.00	17.1
	ค	.19	.41	16.5		ค	.17	.29	16.8
	ง*	.62	.61	11.8		ง*	.37	.32	14.3
	จ	.12	.46	17.6		จ	.14	.05	17.2
25	ก	.16	.60	17.0	30	ก	.19	.27	16.5
	ข	.07	.39	18.8		ข*	.57	.50	12.3
	ค	.05	.39	19.7		ค	.05	.16	19.6
	ง*	.59	.65	12.1		ง	.16	.43	16.9
	จ	.13	.30	17.6		จ	.05	.39	19.7

* เป็นตัวเลือกที่ถูกต้อง

ตาราง 10 (ต่อ)

ข้อ	ตัวเลือก	p_E	r	Δ	ข้อ	ตัวเลือก	p_E	r	Δ
31	ก	.06	.47	19.1	36	ก	.11	.61	117.8
	ข	.05	.39	19.7		ข	.05	.39	19.7
	ค*	.67	.57	11.3		ค	.06	.47	19.1
	ง	.16	.43	16.9		ง*	.61	.70	11.9
	จ	.08	.26	18.5		จ	.14	.28	17.4
32	ก	.10	.26	18.1	37	ก	.05	.39	19.7
	ข	.05	.39	19.7		ข	.11	.18	18.0
	ค	.13	.25	17.5		ค	.21	.34	16.2
	ง	.29	.24	15.3		ง	.24	.17	15.8
	จ*	.46	.44	13.4		จ*	.41	.46	14.0
33	ก	.18	.12	16.7	38	ก	.17	.21	16.8
	ข	.10	.57	18.2		ข	.17	.13	16.8
	ค*	.62	.48	11.8		ค*	.43	.41	13.7
	ง	.05	.39	19.7		ง	.08	.05	18.6
	จ	.06	.47	19.1		จ	.13	.31	17.5
34	ก*	.65	.46	11.4	39	ก	.05	.39	19.7
	ข	.06	.47	19.1		ข	.15	.23	17.1
	ค	.05	.00	19.6		ค	.17	.24	16.8
	ง	.13	.27	17.4		ง	.29	.39	15.2
	จ	.08	.40	18.7		จ*	.31	.69	14.9
35	ก	.05	.39	19.7	40	ก	.05	.39	19.7
	ข	.05	.39	19.7		ข	.11	.10	17.8
	ค	.13	.64	17.5		ค	.15	.19	17.1
	ง	.14	.34	17.3		ง	.39	.07	14.1
	จ*	.65	.67	11.5		จ*	.32	.26	14.9

* เป็นตัวเลือกที่ถูกต้อง

ตาราง 10 (ต่อ)

ข้อ	ตัวเลือก	p _E	r	Δ	ข้อ	ตัวเลือก	p _E	r	Δ
41	ก	.06	.47	19.1	46	ก	.05	.39	19.7
	ข	.10	.24	18.2		ข	.11	.18	18.0
	ค	.05	.39	19.7		ค*	.51	.22	12.9
	ง	.18	.24	16.6		ง	.08	.05	18.6
	จ*	.58	.50	12.2		จ	.29	.14	15.2
42	ก	.13	.24	17.6	47	ก	.22	.00	16.1
	ข	.05	.39	19.7		ข	.21	.17	16.3
	ค*	.47	.37	13.3		ค	.05	.39	19.7
	ง	.15	.19	17.1		ง*	.47	.22	13.3
	จ	.23	.17	15.9		จ	.05	.08	19.6
43	ก	.10	.00	18.1	48	ก	.08	.05	18.6
	ข	.15	.16	17.1		ข	.11	.00	17.9
	ค	.09	.10	18.4		ค*	.53	.20	12.7
	ง*	.39	.29	14.1		ง	.13	.13	17.4
	จ	.26	.16	15.5		จ	.14	.20	17.3
44	ก	.08	.05	18.6	49	ก	.17	.24	16.8
	ข	.14	.16	17.3		ข	.11	.00	17.9
	ค	.14	.15	17.3		ค	.08	.05	18.6
	ง	.05	.39	19.7		ง	.18	.08	16.6
	จ*	.63	.20	11.7		จ*	.44	.22	13.6
45	ก	.11	.13	17.9	50	ก*	.41	.33	14.0
	ข	.12	.21	17.8		ข	.07	.00	18.9
	ค	.13	.36	17.4		ค	.09	.10	18.4
	ง	.22	.37	16.0		ง	.19	.12	16.5
	จ*	.36	.65	14.4		จ	.23	.27	15.9

* เป็นตัวเลือกที่ถูกต้อง

ตาราง 10 (ต่อ)

ข้อ	ตัวเลือก	p_E	r	Δ	ข้อ	ตัวเลือก	p_E	r	Δ
51	ก*	.39	.35	14.1	56	ก	.13	.27	17.4
	ข	.07	.12	19.0		ข	.10	.00	18.1
	ค	.12	.10	17.6		ค	.12	.10	17.6
	ง	.11	.18	18.0		ง*	.63	.20	11.7
	จ	.30	.19	15.1		จ	.05	.39	19.7
52	ก	.10	.00	18.1	57	ก	.08	.05	18.6
	ข	.12	.10	17.6		ข	.11	.04	17.9
	ค*	.69	.20	11.1		ค	.09	.10	18.4
	ง	.05	.39	19.7		ง*	.63	.20	11.7
	จ	.08	.31	18.7		จ	.09	.28	18.4
53	ก	.14	.11	17.3	58	ก	.05	.04	19.4
	ข	.15	.00	17.1		ข	.05	.39	19.7
	ค	.26	.13	15.6		ค	.18	.09	16.7
	ง	.18	.00	16.7		ง	.39	.12	14.1
	จ*	.27	.20	15.5		จ*	.36	.20	14.5
54	ก*	.24	.20	15.8	59	ก	.08	.05	18.6
	ข	.10	.02	18.0		ข	.13	.12	17.5
	ค	.05	.39	19.7		ค*	.68	.20	11.1
	ง	.30	.12	15.1		ง	.05	.39	19.7
	จ	.34	.04	14.7		จ	.10	.24	18.2
55	ก	.08	.08	18.5	60	ก	.10	.00	18.1
	ข*	.32	.22	14.8		ข	.09	.10	18.4
	ค	.12	.12	17.7		ค	.27	.15	15.5
	ง	.28	.28	15.3		ง*	.34	.20	14.7
	จ	.19	.19	16.5		จ	.20	.03	16.4

* เป็นตัวเลือกที่ถูกต้อง

ภาคผนวก ค

ค่าดัชนีความลำเอียงของข้อสอบ

ตาราง 11 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายต่ำ จำนวน 20 ข้อ จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์และวิธีแมนเทิล-แฮนส์เซล

ข้อที่	β_U	p-value	α_{MH}	p-value
3	0.136	0.001*	1.89	0.001*
6	0.008	0.869	0.97	0.926
10	-0.074	0.057	0.77	0.278
14	-0.084	0.071	0.77	0.185
16	-0.044	0.251	0.88	0.633
17	0.046	0.271	1.41	0.088
18	0.024	0.578	1.26	0.293
21	0.101	0.022*	1.63	0.015*
22	0.032	0.482	1.28	0.215
29	0.091	0.013*	2.22	0.000*
39	-0.011	0.790	1.00	0.935
40	-0.048	0.298	0.84	0.418
43	-0.048	0.259	0.78	0.257
45	-0.078	0.048*	0.64	0.049*
51	-0.038	0.297	0.68	0.089
53	0.046	0.308	1.52	0.057
54	-0.035	0.399	0.89	0.587
55	-0.034	0.417	0.92	0.742
58	-0.006	0.876	0.89	0.653
60	0.012	0.768	1.17	0.479

* หมายถึง ข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 12 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายปานกลาง จำนวน 20 ข้อ จากการ
ตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์และวิธีแมนเทิล-แฮนส์เซล

ข้อที่	β_U	p-value	α_{MH}	p-value
7	0.038	0.311	1.25	0.355
8	-0.080	0.073	0.80	0.271
9	-0.031	0.471	0.94	0.804
11	-0.041	0.323	0.81	0.334
13	0.009	0.822	0.97	0.967
15	-0.059	0.186	0.90	0.639
23	0.331	0.000*	5.85	0.000*
25	0.018	0.554	1.31	0.328
27	0.045	0.145	1.50	0.164
30	0.033	0.249	1.20	0.585
32	-0.044	0.283	0.82	0.388
37	-0.023	0.569	0.86	0.491
38	-0.083	0.036*	0.66	0.049*
41	0.051	0.079	0.90	0.787
42	-0.024	0.584	0.97	0.954
46	0.007	0.876	1.03	0.934
47	0.011	0.640	1.36	0.476
48	-0.043	0.298	0.98	0.993
49	-0.112	0.004*	0.55	0.005*
50	-0.073	0.100	0.76	0.168

* หมายถึง ข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 13 ค่าดัชนีความลำเอียงของข้อสอบที่มีระดับความง่ายสูง จำนวน 20 ข้อ จากการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีชิปเทสต์และวิธีแมนเทล-แฮนส์เซล

ข้อที่	β_U	p-value	α_{MH}	p-value
1	0.058	0.143	1.67	0.027*
2	-0.061	0.135	0.80	0.262
4	-0.035	0.368	0.88	0.615
5	0.013	0.614	1.20	0.590
12	-0.047	0.312	0.93	0.775
19	-0.067	0.143	0.85	0.422
20	-0.079	0.080	0.77	0.185
24	0.034	0.299	1.43	0.163
26	0.011	0.622	0.60	0.262
28	-0.014	0.512	0.81	0.685
31	-0.030	0.440	0.87	0.570
33	-0.073	0.115	0.74	0.121
34	0.045	0.232	1.49	0.068
35	-0.022	0.426	0.61	0.146
36	-0.008	0.754	1.14	0.777
44	-0.061	0.145	0.77	0.211
52	-0.012	0.644	0.89	0.770
56	-0.059	0.143	0.81	0.355
57	-0.036	0.339	0.77	0.249
59	0.019	0.588	1.22	0.423

* ข้อสอบที่มีความลำเอียงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ภาคผนวก ง

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบเครื่องมือวิจัย

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบเครื่องมือวิจัย

รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
รองศาสตราจารย์นิภา ศรีไพโรจน์	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
รองศาสตราจารย์วิญญา วิศาลาภรณ์	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์ชวลิต รวยอาจิน	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ - ชื่อสกุล	นางสาววิภา จำนัณ
วัน เดือน ปีเกิด	26 กันยายน 2516
สถานที่เกิด	อำเภอพัฒนานิคม จังหวัดลพบุรี
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 31 หมู่ 4 ตำบลมะนาวหวาน อำเภอพัฒนานิคม จังหวัดลพบุรี 15140
ตำแหน่งหน้าที่การงานในปัจจุบัน	อาจารย์ 1 ระดับ 4
สถานที่ทำงานปัจจุบัน	โรงเรียนวัดสุนทรเทพคีรี อำเภอชัยบาดาล สำนักงานการประถมศึกษาจังหวัดลพบุรี
ประวัติการศึกษา	
พ.ศ. 2532	ชั้นมัธยมศึกษาตอนต้น จากโรงเรียนพัฒนานิคม
พ.ศ. 2535	ชั้นมัธยมศึกษาตอนปลาย จากโรงเรียนพิบูลวิทยาลัย
พ.ศ. 2539	ค.บ. (เกียรตินิยมอันดับ 1) วิชาเอกคณิตศาสตร์, วิชาโทคอมพิวเตอร์ ในโครงการครูทายาท จากสถาบันราชภัฏพระนครศรีอยุธยา
พ.ศ. 2544	กศ.ม. (วิชาเอกการวัดผลการศึกษา) จากมหาวิทยาลัย ศรีนครินทรวิโรฒ