

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2545

การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนนิเวศวิทยาศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลล์ที่แตกต่างกัน

บทคัดย่อ
ของ
กัญญรัตน์ ทรงประโคน

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2545

กัญญรัตน์ ทรงประโคน.(2545).การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกัน.
ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์ ดร. บุญเชิด
บุญโญนนตพงษ์, รองศาสตราจารย์ นิภา ศรีไพโรจน์.

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
วิชาวิทยาศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันสองแบบ คือ แบบที่ 1 กฎชิลลิ่งแบบคะแนนศูนย์
ต่อหนึ่ง และแบบที่ 2 กฎชิลลิ่งแบบคะแนนศูนย์ต่อหนึ่งมากกว่า 1 ตัว และพิจารณาจากการตอบผิด
ตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด โดยมีการจัดเรียง
ข้อสอบตามค่าความยากง่าย

กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2544
ของโรงเรียนในสังกัดกรุงเทพมหานคร จำนวน 438 คน ได้มาโดยการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้
ในการศึกษาเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ผลการวิจัยสรุปได้ดังนี้

1. การให้คะแนนตามกฎชิลลิ่งแบบที่ 1 ความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบวัดผลสัมฤทธิ์
ทางการเรียนวิชาวิทยาศาสตร์ ที่ตรวจจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามและ
การไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยความเชื่อมั่น
ของการตอบผิดตำแหน่งที่หนึ่งมีค่าสูงสุดเท่ากับ .907

2. การให้คะแนนตามกฎชิลลิ่งแบบที่ 2 ความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบวัดผลสัมฤทธิ์
ทางการเรียนวิชาวิทยาศาสตร์ ที่ตรวจจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามและ
การไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยความเชื่อมั่น
ของการตอบผิดตำแหน่งที่สองมีค่าสูงสุดเท่ากับ .901

3. การตรวจให้คะแนนจากการตอบผิดตำแหน่งที่หนึ่ง ความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์
ทางการเรียนวิชาวิทยาศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันสองแบบมีค่าแตกต่างกันอย่างไม่มี
นัยสำคัญทางสถิติ

4. การตรวจให้คะแนนจากการตอบผิดตำแหน่งที่สอง ความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์
ทางการเรียนวิชาวิทยาศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันสองแบบมีค่าแตกต่างกันอย่างไม่มี
นัยสำคัญทางสถิติ

5. การตรวจให้คะแนนจากการตอบผิดตำแหน่งที่สาม ความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์
ทางการเรียนวิชาวิทยาศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันสองแบบมีค่าแตกต่างกันอย่างไม่มี
นัยสำคัญทางสถิติ

THE EFFECT OF DIFFERENT CEILING RULES ON THE INTERNAL
CONSISTENCY RELIABILITY OF SCIENCE ACHIEVEMENT TEST
IN PRATHOM SUKSA VI STUDENTS

AN ABSTRACT
BY
KANYARAT SONGPRAKHON

Presented in partial fulfillment of the requirements
for the Master of Educational Measurement
at Srinakharinwirot University
May 2002

Kanyarat Songprakhon. (2002). *The Effect of Different Ceiling Rules on the Internal Consistency Reliability of Science Achievement Test in Prathom Suksa VI Students*. Master thesis, M.Ed. (Educational Measurement). Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee : Assoc. Prof. Dr. Boonchird Pinyoanuntapong, Assoc. Prof. Nipa Sripairot.

The purpose of this research was to study the effects on the reliability of the science achievement test scored by using different ceiling rules. Type 1, the ceiling rule was determined according to one, two and three consecutive failures found. Type 2, the ceiling rule determined according to more than one consecutive failure found within 3 times consecutively. The unlimited failure or no ceiling was also considered as a control group within two types of different ceiling rules. The test were arranged in order of difficulty.

The sample was 438 Prathom Suksa VI Students in Bangkok Province selected by stratified random sampling technique. The instrument used in this study was a science achievement test.

The results were as follows :

1. The reliability of the science achievement test scored by using ceiling rule type 1 was different significantly at .01 level. The first consecutive failure found yielded the highest reliability at .907
2. The reliability of the science achievement test scored by using ceiling rule type 2 was different significantly at .01 level. The second consecutive failure found yielded the highest reliability at .901
3. The reliability of the science achievement test of two different ceiling rules , for the first consecutive failure found , was not different significantly.
4. The reliability of the science achievement test of two different ceiling rules , for the second consecutive failure found , was not different significantly.
5. The reliability of the science achievement test of two different ceiling rules , for the third consecutive failure found , was not different significantly.

ปริญญานิพนธ์
เรื่อง

การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนนิชาวิทยาศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิ่งที่แตกต่างกัน

ของ
นางสาวกัญญรัตน์ ทรงประโคน

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

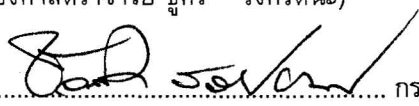

..... คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร. นภาพร หะวานนท์)
วันที่ ๘ เดือน พฤษภาคม พ.ศ. 2545

คณะกรรมการสอบปริญญานิพนธ์


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


..... กรรมการ
(รองศาสตราจารย์ นิภา ศรีไพโรจน์)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ สุศรี วงศ์รัตนะ)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ชวลิต รวยอาจิณ)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี โดยได้รับความอนุเคราะห์ให้คำปรึกษาชี้แนะแนวทาง และแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดียิ่ง จาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์ รองศาสตราจารย์นิภา ศรีไพโรจน์ ตลอดจนคณาจารย์ภาควิชาการวัดผลและวิจัยการศึกษาทุกท่านที่ได้กรุณาให้ความคิดเห็น ข้อเสนอ ซึ่งผู้วิจัยขอกราบขอบพระคุณมา ณ ที่นี้

ขอกราบขอบพระคุณ รองศาสตราจารย์ชูศรี วงศ์รัตนะ และอาจารย์ชวลิต รวยอาจิณที่กรุณาเป็นกรรมการสอบปากเปล่าปริญญานิพนธ์ ที่ให้แนวคิด และชี้แนะข้อบกพร่อง

ในการตรวจและพิจารณาเครื่องมือที่ใช้ในการวิจัย ได้รับความเอื้อเฟื้อจาก ผู้ช่วยศาสตราจารย์ ระวีวรรณ พันธุ์พานิช อาจารย์ชวลิต รวยอาจิณ อาจารย์อุดมศักดิ์ นาคี อาจารย์จินลักษณะ ธรรมทวีวิกุล และอาจารย์วรรณ พิทักษ์ถิร ซึ่งผู้วิจัยขอขอบพระคุณเป็นอย่างสูง

ขอขอบคุณอาจารย์มิ่ง เทพครเมือง และอาจารย์พิมพ์ ภูมิรา ที่ให้ความรู้เกี่ยวกับวิธีการตรวจ ให้คะแนนโดยกฎซิกลิ่ง ตลอดจนคำแนะนำในการวิเคราะห์ข้อมูลทางสถิติ

การเก็บรวบรวมข้อมูลได้รับความร่วมมือจากผู้บริหาร คณะครูและนักเรียนของโรงเรียนสังกัด กรุงเทพมหานคร สำนักงานเขตตลิ่งชันที่เป็นกลุ่มตัวอย่าง รวมทั้งอาจารย์จุฑาดี ทาхла ที่ให้การสนับสนุน และร่วมมือช่วยเหลือในการเก็บรวบรวมข้อมูล ซึ่งผู้วิจัยรู้สึกซาบซึ้งในความกรุณาที่ทุกท่านให้ความช่วยเหลือเป็นอย่างดี

ขอขอบคุณเพื่อนครูจากโรงเรียนราชินีทุกท่าน เพื่อน ๆ วัดผล ภาคพิเศษรุ่น 9 ที่สนับสนุนและให้ความช่วยเหลือ ตลอดจนเป็นกำลังใจในการทำปริญญานิพนธ์ฉบับนี้ และขอขอบคุณคุณจุฑา ทรวงประโคน และคุณพัชรี ทรวงประโคน สำหรับความช่วยเหลือและเป็นแรงผลักดันในการทำงานจนสำเร็จ คุณค่าและสาระประโยชน์ของปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอบแต่ บิดา – มารดา ผู้เป็นกำลังใจอันยิ่งใหญ่และ ผู้ให้โอกาสทางการศึกษา

กัญญรัตน์ ทรวงประโคน

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการศึกษาค้นคว้า.....	4
ความสำคัญของการศึกษาค้นคว้า.....	4
ขอบเขตของการศึกษาค้นคว้า.....	5
ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย.....	5
ตัวแปรที่ใช้ในการศึกษาค้นคว้า.....	5
นิยามศัพท์เฉพาะ.....	6
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	9
วิธีการตรวจให้คะแนน.....	9
การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้.....	15
การเรียงข้อคำถามในแบบทดสอบ.....	20
ความเที่ยงตรงของแบบทดสอบ.....	22
ความเชื่อมั่นของแบบทดสอบ.....	26
เอกสารที่เกี่ยวข้องกับการสอนวิทยาศาสตร์ในโรงเรียนประถมศึกษา.....	44
งานวิจัยที่เกี่ยวข้องกับวิธีการตรวจให้คะแนน.....	46
งานวิจัยที่เกี่ยวข้องกับการเรียงข้อคำถามในแบบทดสอบ.....	54
สมมติฐานในการศึกษาค้นคว้า.....	60
3 วิธีดำเนินการศึกษาค้นคว้า.....	61
ประชากร.....	61
กลุ่มตัวอย่าง.....	61
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	62
วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	63
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	69
วิธีการตรวจให้คะแนนโดยกฎซิงสองแบบ.....	69
ขั้นตอนในการวิเคราะห์ข้อมูล.....	76
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	72
4 ผลการวิเคราะห์ข้อมูล.....	75
สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล.....	75
ผลการวิเคราะห์ข้อมูล.....	75

สารบัญ(ต่อ)

บทที่	หน้า
5 สรุป อภิปรายผล และข้อเสนอแนะ.....	82
ความมุ่งหมาย สมมติฐาน และเครื่องมือที่ใช้ในการเก็บข้อมูล.....	83
สรุปผลการศึกษาค้นคว้า.....	84
อภิปรายผล.....	85
ข้อเสนอแนะ.....	87
 บรรณานุกรม.....	 88
 ภาคผนวก.....	 97
ภาคผนวก ก.....	98
ภาคผนวก ข.....	100
ภาคผนวก ค.....	102
ภาคผนวก ง.....	106
 ประวัติย่อผู้วิจัย.....	 112

บัญชีตาราง

ตาราง	หน้า
1 วิธีการให้คะแนนแบบเลือกเสรี.....	10
2 วิธีให้คะแนนแบบบอกระดับความมั่นใจ.....	10
3 วิธีให้คะแนนแบบกิบบอนส์ โอลส์คิน และโซเบล.....	11
4 วิธีการตรวจให้คะแนนของเกริกชัย ฮวบเจริญ.....	12
5 วิธีการตรวจให้คะแนนของสำราญ มีแจ้ง.....	13
6 เมทริกซ์ของความเที่ยงตรงแบบ Multitrait Multimethod.....	25
7 เมทริกซ์ของความเที่ยงตรงแบบ Multitrait Multimethod ที่วัด 2 ลักษณะ (Traits) และใช้วิธีวัด 2 วิธี (Methods).....	26
8 แสดงหน่วยการเรียนรู้และหน่วยการเรียนรู้ย่อย กลุ่มสร้างเสริมประสบการณ์ชีวิต จำแนกเป็นระดับชั้น.....	46
9 จำนวนนักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัย จำแนกตามขนาดโรงเรียน.....	62
10 ตารางวิเคราะห์รายละเอียดวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6.....	65
11 การวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน ทั้งสองแบบ.....	76
12 ค่าความเชื่อมั่น ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบ จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด.....	77
13 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนน โดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด.....	78
14 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิง แบบศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนด ตำแหน่งการตอบผิด.....	79
15 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดย กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด.....	79
16 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิง แบบศูนย์ต่อเนื่องมากกว่า 1 (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด.....	80

บัญชีตาราง(ต่อ)

ตาราง	หน้า
17 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซัลลิ่งที่ แตกต่างกันทั้ง 2 แบบของแบบทดสอบ จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม.....	81
18 ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้.....	101
19 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดผลสัมฤทธิ์ ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 จากการทดลองใช้(Try out).....	103
20 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดผลสัมฤทธิ์ ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ใช้ในการวิจัย.....	105

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 ลำดับขั้นตอนในการดำเนินการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามพฤติกรรมการเรียนรู้ของ คลอปเฟอร์.....	63

บทที่ 1

บทนำ

ภูมิหลัง

การศึกษาเป็นรากฐานที่สำคัญยิ่งต่อการพัฒนาประเทศในทุก ๆ ด้าน โดยเฉพาะการพัฒนาทรัพยากรมนุษย์ให้มีความเจริญงอกงาม มีคุณสมบัติตรงตามความต้องการของสังคม สามารถดำรงชีวิตอย่างมีความสุข รวมทั้งมีความสามารถในการประกอบอาชีพและเป็นกำลังในการพัฒนาประเทศให้ทัดเทียมนานาอารยประเทศ เมื่อการศึกษามีแนวโน้มไปในทางพัฒนากำลังคนของชาติ การจัดการศึกษาจึงควรจัดให้มีประสิทธิภาพตรงตามวัตถุประสงค์ที่วางไว้ การวัดผลการศึกษาจึงเป็นขั้นตอนสำคัญที่จะตรวจสอบผลการศึกษาว่า ผลการศึกษาด้วยวิธีการต่างๆ มีผลสัมฤทธิ์มากน้อยเพียงใด เพื่อจะได้นำผลที่ได้จากการวัดนั้นมาใช้ในการปรับปรุงการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น จึงถือว่าการวัดผลประเมินผลเพื่อพัฒนาการเรียนการสอน ดังนั้น การที่จะจัดการศึกษาให้บรรลุจุดหมายได้อย่างมีประสิทธิภาพต้องประกอบไปด้วยองค์ประกอบ 5 อย่าง คือ ปรัชญาการศึกษา หลักสูตร การเรียนการสอน การวัดผลการศึกษา และการวิจัย (ต่าย เชียงฉี. 2526 : 1) การวัดผลการศึกษาเป็นองค์ประกอบที่สำคัญประการหนึ่งของกระบวนการเรียนการสอนในระดับการศึกษาต่างๆ เพราะการวัดผลการศึกษาเป็นเครื่องมืออันหนึ่งที่จะช่วยพัฒนาคุณภาพการศึกษา ผลจากการวัดจะเป็นพื้นฐานในการตัดสินใจของครู เพื่อใช้ปรับปรุงวิธีการสอน การแนะแนว การประเมินผล หลักสูตร แบบเรียน การใช้อุปกรณ์การสอน ทั้งยังปรับปรุงการเรียนของนักเรียนให้ก้าวหน้ายิ่งขึ้น และผลการเรียนที่ไม่ดีของนักเรียนไม่เพียงแสดงถึงความอ่อนแอของนักเรียนแต่ละบุคคลเท่านั้น แต่ถ้าพิจารณาผลการสอนรวมทั้งชั้นก็จะแสดงถึงผลการสอนที่ไม่ดีของครูได้อีกด้วย (อนันต์ ศรีโสภา. 2520 : 1) ทั้งนี้เนื่องจากการวัดผลการศึกษาจะเป็นตัวชี้ให้เห็นว่ากระบวนการให้การศึกษานั้นสัมฤทธิ์ผลเพียงใด การวัดผลการศึกษานั้นสามารถกระทำได้หลายวิธี แต่ที่นิยมกันมากก็คือ การวัดผลโดยใช้แบบทดสอบ ซึ่งการวัดผลโดยใช้แบบทดสอบจะมีประสิทธิภาพเพียงใดขึ้นอยู่กับคุณภาพของแบบทดสอบที่ใช้โดยตรงเพราะถ้าแบบทดสอบที่สร้างขึ้น เป็นแบบทดสอบที่ขาดคุณลักษณะที่สำคัญของแบบทดสอบที่ดี ก็ไม่สามารถสัมฤทธิ์ผลได้ นักวัดผลและนักการศึกษาได้กล่าวถึงคุณลักษณะหรือคุณสมบัติของแบบทดสอบไว้ว่าคุณลักษณะสำคัญของแบบทดสอบที่ดีมี 3 ประการ คือ ความเที่ยงตรง ความเชื่อมั่น และสามารถนำไปใช้ได้ (Lyman. 1963 : 25, Chase. 1974 : 56)

สิ่งสำคัญที่สุดในการทดสอบก็คือ แบบทดสอบซึ่งในการวัดผลสัมฤทธิ์ทางการเรียนนั้น ทั้งชาวแพร์ตกุล (ชวาล แพร์ตกุล. 2518 : 164) อีเบล (Ebel. 1965 : 60) และธอร์นไดค์ (Thorndike. 1969 : 58) มีความเห็นสอดคล้องกันว่า แบบทดสอบปรนัยชนิดเลือกตอบเป็นแบบทดสอบที่ดีที่สุด เพราะแบบทดสอบปรนัยชนิดเลือกตอบสามารถเขียนคำถามได้ครอบคลุมเนื้อหาและวัตถุประสงค์ มีความเป็นปรนัยในการตรวจให้คะแนน และมีประสิทธิภาพ ดังที่ เรมเมอร์ (Remmer. 1955 : 94-95) ได้กล่าวว่า แบบทดสอบปรนัยชนิดเลือกตอบมีข้อดีที่วัดเนื้อหาได้ครอบคลุม ให้คะแนนยุติธรรม ส่วนอาดัมส์ (Adams. 1964 : 331) กล่าวว่าไว้ว่า แบบทดสอบปรนัยชนิดเลือกตอบวัดความจำเกี่ยวกับข้อเท็จจริงได้ดี และสามารถสร้างให้วัดสมรรถภาพสมองด้านต่างๆ ได้ ซึ่งเหมาะที่จะใช้ทดสอบเมื่อมีผู้สอบจำนวนมาก

แม้ว่าแบบทดสอบปรนัยแบบเลือกตอบจะมีข้ออยู่หลายประการแต่นักการศึกษาหลายท่านเห็นว่า มีข้อเสียด้านส่งเสริมการเดา (พจน์ สะเพียรชัย. 2515 : 21) ทำให้ไม่สามารถแยกผู้ที่มีความรู้ได้ถูกต้อง ลิทเทิล (วัชรภรณ์ จิตรมาศ. 2538 : 2 ; อ้างอิงจาก Little 1962) กล่าวถึงลักษณะระดับความรู้ของผู้ตอบที่นำไปสู่การตอบถูกหรือผิดเกี่ยวข้องกัน 6 ลักษณะ คือ มีความรู้ถูกต้องในเรื่องนั้นนำไปสู่การตอบที่ถูก มีความรู้บางส่วนนำไปสู่การตอบที่ถูก มีความรู้บางส่วนนำไปสู่การตอบที่ผิด ไม่มีความรู้ทำให้เลือกตอบแบบเดาสุ่ม มีความรู้ไม่ถูกต้องนำไปสู่การตอบผิด รู้สึกว่าคำถามในข้อสอบไม่เหมาะสม รู้คำตอบที่ผู้วัดต้องการให้ตอบแต่ไม่เห็นด้วยจึงไม่เลือกคำตอบที่ถูกต้อง นอกจากนี้ในการตอบข้อสอบนักเรียนจะพิจารณาหาข้อที่ถูกที่สุดเพียงข้อเดียวเมื่อพบแล้วจะไม่พิจารณาตัวเลือกอื่น ๆ โดยจะถือว่าตัวเลือกที่เลือกผิดทันที ถ้านักเรียนไม่ทราบมาก่อนว่าทำไมจึงผิด หรืออาจตอบโดยค่อย ๆ กำจัดตัวเลือกที่ผิดออกไป ซึ่งทำได้ง่าย เพราะโอกาสที่เลือกตัวเลือกตอบต่าง ๆ เป็นตัวเลือกที่ผิดนั้นมีมาก ทำให้ผลจากการวัดเป็นคะแนนจริงรวมกับคะแนนความคลาดเคลื่อนไม่ใช่ความสามารถที่แท้จริงของผู้เรียนจึงต้องมีการคิดวิธีแก้การเดาหลายรูปแบบ เช่น สร้างข้อสอบให้มีตัวเลือกเป็นเอกพันธ์ สร้างจำนวนตัวเลือกหลายๆ ตัว หรือใช้สูตรแก้การเดา

ปัญหาเรื่องการเดามีผลต่อค่าความเชื่อมั่นและคะแนนความคลาดเคลื่อน ถ้าสามารถกำหนดเงื่อนไขในการให้นักเรียนตอบและเงื่อนไขการตรวจให้คะแนนได้อย่างถูกต้องเหมาะสมจะช่วยให้คะแนนที่ได้มีค่าสหสัมพันธ์สูงกับความสามารถที่แท้จริงของนักเรียน (อรรพรรณ ตันท์เจริญรัตน์. 2517 : 2 - 3)

โรเบิร์ต แอล ธอร์นไดส์ (Thorndike. 1971 : 60) เสนอความคิดเห็นว่า การขจัดปัญหาเกี่ยวกับการเดาสามารถทำได้ 2 วิธี คือ การปรับเทคนิคการให้คะแนนหลังจากสอบเสร็จแล้ว และวิธีการให้คำสั่งชี้แจงก่อนทำข้อสอบ วิธีการดังกล่าว เช่น วิธีของคัมบ์ส (Coombs. 1956 : 308 - 313) ซึ่งเสนอให้เลือกคำตอบที่ผิด แทนการเลือกคำตอบที่ถูก วิธีของเดวิด และฟีฟเฟอร์ (Davis and Fiffer. 1953 : 159 - 169)

ซึ่งเสนอการตรวจให้คะแนนโดยกำหนดให้น้ำหนักของคะแนนของแต่ละตัวเลือกไม่เท่ากัน นอกจากนี้สมบุรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) วิลสันและคณะ (Wilson, et al. 1988 : 213 - 217) เสนอวิธีการตรวจให้คะแนนโดยกฎซิลลิง (Ceiling Rules) ที่เรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดตามรูปแบบของกัตแมน (Guttman Model) โดยถือว่าคำตอบข้อที่นักเรียนตอบถูกที่อยู่หลังคำตอบข้อที่นักเรียนตอบผิดเป็นการตอบถูกด้วยการเดามากกว่าเป็นการตอบจากความสามารถที่แท้จริง การตอบผิดของแต่ละบุคคลจะมีระดับต่าง ๆ กัน ขึ้นอยู่กับความสามารถของผู้ตอบและระดับความยากของแบบทดสอบที่ใช้

มีงานวิจัยในต่างประเทศเกี่ยวกับวิธีการตรวจให้คะแนนโดยกฎซิลลิง สมบุรณ์ สุริยวงศ์

(Suriyawongse, Somboon. 1987 : 613 - A) ศึกษาเปรียบเทียบค่าความเชื่อมั่นแบบทดสอบมาตรฐานวิชาคณิตศาสตร์กับนักเรียนเกรด 8 โดยเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดแล้วตรวจให้คะแนนโดยกฎซิลลิงจากการกำหนดตำแหน่งการตอบผิด (Ceiling Point) ต่างกัน 5 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด เพื่อหาความสามารถที่แท้จริงโดยให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด (0) เป็นข้อแรก จึงหยุดตรวจแล้วนับคะแนน โดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่หนึ่งคำตอบถูกหลังการตอบผิดตำแหน่งที่หนึ่งจะไม่นำมานับคะแนน และให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรกจึงหยุดตรวจแล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดติดต่อกันนั้นเป็นการตอบผิดตำแหน่งที่สอง คำตอบถูกหลังการตอบผิดตำแหน่งที่สองจะไม่นำมานับคะแนน ในทำนองเดียวกันตำแหน่งการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าจะให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สามข้อ (000) สี่ข้อ (0000) และห้าข้อ (00000) เป็นครั้งแรกจึงหยุดตรวจแล้วนับคะแนน โดยถือว่าข้อที่ตอบผิดนั้น ๆ เป็นการตอบผิดตำแหน่งที่สาม

ตำแหน่งที่สี่ และตำแหน่งที่ห้าตามลำดับ ส่วนคำตอบถูกหลังการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้า จะไม่นำนับคะแนน สำหรับการไม่กำหนดตำแหน่งการตอบผิดจะให้ข้อละ 1 คะแนนกับทุกข้อที่ตอบถูก ผลที่ได้จากการศึกษาวิจัยพบว่าการตอบผิดตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่ง และการไม่กำหนดตำแหน่งการตอบผิดโดยการตอบผิดตำแหน่งที่สองให้ค่าความเชื่อมั่นสูงสุด

นอกจากวิลสัน และคณะ (Wilson, et al. 1988 : 213 - 216) ได้ศึกษาวิจัยค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์กับนักเรียนเกรด 8 โดยเรียงข้อสอบ จากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด แล้วตรวจให้คะแนนโดยกฎชิลลิ่งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด เพื่อหาความสามารถที่แท้จริงโดยให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด(0) เป็นข้อแรกจึงหยุดตรวจแล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่หนึ่ง คำตอบถูกหลังการตอบผิดตำแหน่งที่หนึ่งจะไม่นำนับคะแนน และให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรกจึงหยุดตรวจ แล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่สอง คำตอบถูกหลังการตอบผิดตำแหน่งที่สองจะไม่นำนับคะแนน ในทำนองเดียวกันการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่และตำแหน่งที่ห้าจะให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สามข้อ (000) สี่ข้อ (0000) และห้าข้อ (00000) เป็นครั้งแรกจึงหยุดตรวจ แล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าตามลำดับ คำตอบถูกหลังการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าจะไม่นำนับคะแนน สำหรับการไม่กำหนดตำแหน่งการตอบผิดจะให้ข้อละ 1 คะแนนกับทุกข้อที่ตอบถูก ผลการศึกษาวิจัยที่ได้สอดคล้องกับ สมบูรณ์ สุริยวงศ์ คือ การตอบผิดตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่ง และการไม่กำหนดตำแหน่งการตอบผิดโดยการตอบผิดตำแหน่งที่สองให้ค่าความเชื่อมั่นสูงสุด

บลูมและคณะ (ลัวัน สายยศ และอังคณา สายยศ. 2539 : 41 ; อ้างอิงจาก Bloom, et al. n.d.) ได้เสนอแนวคิดในการวัดพฤติกรรมการเรียนรู้ด้านพุทธิพิสัย (Cognitive Domain) ไว้ 6 ด้าน ประกอบด้วย ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า โดยเรียงลำดับพฤติกรรมขั้นง่ายไปสู่พฤติกรรมขั้นยาก ซึ่งคลอปเฟอร์ (Leopold E. Klopfer) ได้ศึกษาวัตถุประสงค์ทางการศึกษาของบลูม แล้วนำมากำหนดเป็นวัตถุประสงค์ให้เหมาะสมกับการเรียนการสอนวิชาวิทยาศาสตร์ ซึ่งประกอบด้วย ความจำ ความเข้าใจ กระบวนการสืบเสาะหาความรู้วิทยาศาสตร์ และการนำไปใช้ ดังนั้น การเขียนข้อสอบตามพฤติกรรมการเรียนรู้จึงสอดคล้องกับแบบทดสอบที่มีลักษณะเป็นแบบโมเดลของกัตแมน คือเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด การเรียงข้อสอบลักษณะนี้ใช้กับวิธีการตรวจให้คะแนนโดยกฎชิลลิ่ง

ในปัจจุบันวิธีการตรวจให้คะแนนแบบทดสอบที่ใช้กันอยู่เป็นการให้คะแนนแบบ 0 - 1 คือ ให้ 1 คะแนนสำหรับคำตอบถูก และ 0 คะแนนสำหรับคำตอบผิด ไม่มีการกำหนดตำแหน่งการตอบผิดผู้วิจัยจึงมีความเห็นว่าวิธีการตรวจให้คะแนนเช่นนี้ เป็นวิธีการที่น่าจะมีความคลาดเคลื่อนมากเพราะเป็นการปฏิเสธคะแนนที่เป็นจริง (Type I error) และยอมรับคะแนนที่เป็นเท็จ (Type II error) เข้ามาเป็นคะแนนรวม อย่างไรก็ตาม วิธีการตรวจให้คะแนนโดยกฎชิลลิ่งของ สมบูรณ์ สุริยวงศ์ วิลสัน และคณะ ซึ่งค่อนข้างจะเป็นวิธีการที่เข้มงวดมากเกินไปจึงเป็นการให้อโอกาสการตอบผิดที่ต่อเนื่องกันเพียงครั้งเดียวเป็นจุดตัด ผู้วิจัยเห็นว่าน่าจะเปิดโอกาสการตอบผิดที่ต่อเนื่องกันหลาย ๆ ครั้งเป็นจุดตัดเพื่อหาความสามารถที่เป็นคะแนนจริง

ได้มากขึ้น จากแนวคิดดังกล่าว ผู้วิจัยจึงมีความคิดที่จะศึกษาคำความเชื่อมั่นของแบบทดสอบ จากผลการใช้ กฎซิลลิงของ สมบูรณ์ สุริยวงศ์ วิลสัน และ คณะ (กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง) กับกฎที่เปิดโอกาส การตอบผิดที่ต่อเนื่องกันได้มากกว่า 1 ครั้งติดต่อกัน (กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว) ของ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533)ว่าจะให้ค่าความเชื่อมั่นต่างกันหรือไม่

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาค้นคว้าในครั้งนี้มีจุดมุ่งหมายหลัก

1. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง
2. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว
3. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการ ตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
4. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง มากกว่า 1 ตัว จากการ กำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดตำแหน่งที่หนึ่ง
6. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดตำแหน่งที่สอง
7. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดตำแหน่งที่สาม

ความสำคัญของการศึกษาค้นคว้า

ผลการศึกษาค้นคว้าทำให้ทราบว่า การตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันสองแบบ แบบใด ตำแหน่งใด ส่งผลต่อค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ สูงสุด อันจะเป็นแนวทางในการตรวจให้คะแนน และคะแนนที่ได้จากการวัดที่น่าเชื่อถือที่สุด

ขอบเขตของการศึกษาค้นคว้า

ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย

ประชากรสำหรับสุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 16 โรงเรียน

การเลือกกลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 8 โรงเรียน จำนวน 438 คน โดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling)

ประมาณกลุ่มตัวอย่างที่ใช้ในการศึกษาด้วยความเชื่อมั่น 95% ($\alpha = .05$) เมื่อเทียบจากตารางขนาดของกลุ่มตัวอย่างที่ควรเลือกจากประชากรพบว่าต้องใช้กลุ่มตัวอย่างอย่างน้อย 286 คน (ล้วน สายยศ และอังคณา สายยศ . 2540 : 400 , อ้างอิงจาก Yamane.1967) แต่ในการวิจัยครั้งนี้ใช้กลุ่มตัวอย่าง 438 คน ซึ่งทำการสุ่มดังนี้ แบ่งโรงเรียนออกเป็นชั้นโดยใช้ขนาดโรงเรียนเป็นชั้น (Strata) และโรงเรียนเป็นหน่วยการสุ่ม โดยแบ่งโรงเรียนออกเป็น 3 ขนาด คือ โรงเรียนขนาดใหญ่ โรงเรียนขนาดกลาง และโรงเรียนขนาดเล็ก สุ่มโรงเรียนมาประมาณ 50% ของแต่ละขนาด ได้โรงเรียนขนาดใหญ่ 2 โรงเรียน โรงเรียนขนาดกลาง 4 โรงเรียน และโรงเรียนขนาดเล็ก 2 โรงเรียน

ตัวแปรที่ใช้ในการศึกษาค้นคว้า

1. ตัวแปรอิสระ

1.1 วิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน 2 แบบ

1.1.1 วิธีการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง

1.1.2 วิธีการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

1.2 ตำแหน่งการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด

2. ตัวแปรตาม คือ ความเชื่อมั่นของแบบทดสอบ

เนื้อหาวิชาที่ใช้ในการสร้างแบบทดสอบ

เป็นเนื้อหาวิชาสร้างเสริมประสบการณ์ชีวิต (วิทยาศาสตร์) ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ประกอบด้วยหัวข้อดังต่อไปนี้

1. สสารและความร้อน
2. แสง
3. ไฟฟ้า
4. แรง แรงดัน ความกดดัน
5. สารเคมี

ข้อตกลงเบื้องต้น

1. ในการศึกษาครั้งนี้ นักเรียนต้องทำข้อสอบในแบบทดสอบทุกข้อ และจะนำเฉพาะข้อมูลที่มีการตอบผิดถึงตำแหน่งที่ 3 ของทั้งสองแบบมาวิเคราะห์
2. ข้อสอบที่นักเรียนตอบผิดตั้งแต่ข้อแรกจะไม่นำมาวิเคราะห์
3. ข้อสอบที่นักเรียนทำถูก หรือ ผิดหมดทุกข้อจะไม่นำมาวิเคราะห์

นิยามศัพท์เฉพาะ

1. จุดซัลลิ่ง หรือ ตำแหน่งการตอบผิด (Ceiling Point)

หมายถึง จุดหรือตำแหน่งของข้อสอบข้อที่นักเรียนตอบผิดเป็นตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

2. กฎซัลลิ่ง (Ceiling Rules)

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ข้อสอบข้อที่นักเรียนตอบถูกหลังข้อที่ตอบผิดมีค่าความยากมากกว่า ถือว่านักเรียนตอบข้อสอบข้อนั้นโดยมีแนวโน้มที่จะเดามากกว่าใช้ความรู้จริง การศึกษาครั้งนี้มี

- 2.1 กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่อง

- 2.2 กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

3. กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่อง

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิดเป็นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ซึ่งมีการกำหนดคะแนนในแต่ละตำแหน่ง ดังนี้

3.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 – 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิด (0) เป็นข้อแรก ข้อที่ตอบถูกหลังจากนั้นไม่นำมานับคะแนน

3.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 – 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันสองข้อ(00)เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันสามข้อ(000)เป็นครั้งแรกข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4. กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิดเป็นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิดเมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ซึ่งมีการกำหนดคะแนนในแต่ละตำแหน่ง ดังนี้

4.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สอง ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สาม ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

5. การไม่กำหนดตำแหน่งการตอบผิด

หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจแบบปกติ ให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

6. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

หมายถึง แบบทดสอบที่ผู้วิจัยสร้างขึ้น ประกอบด้วยชุดของข้อคำถามแบบปรนัยเลือกตอบชนิด 4 ตัวเลือก วิชาวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 6 เนื้อหาเรื่อง สสารและความร้อน แสง ไฟฟ้า แรง แรงดัน ความกดดัน และสารเคมี ตามพฤติกรรมการเรียนรู้ของคลอปเฟอร์ ซึ่งใช้ระดับพฤติกรรมการเรียนรู้

4 พฤติกรรม คือ ความรู้ความจำ ความเข้าใจ กระบวนการสืบเสาะหาความรู้วิทยาศาสตร์ และการนำไปใช้

7. การจัดเรียงข้อสอบตามความยากง่าย

หมายถึง การเรียงข้อสอบจากข้อที่ง่ายไปยังข้อที่ยาก (ค่า p อยู่ระหว่าง .80 - .20)

8. ความเชื่อมั่นของแบบทดสอบ

หมายถึง คุณสมบัติของแบบทดสอบที่สามารถวัดความรู้ ความสามารถของนักเรียนได้คงที่แน่นอน หรือความคงเส้นคงวาของคะแนนจากการสอบวัด ในการศึกษาครั้งนี้หาค่าความเชื่อมั่น โดยใช้สูตรสัมประสิทธิ์ r_B (Coefficient r_B) (บุญเชิด ภิญญอนันตพงษ์. 2545 : 69)

9. หลักฐานแสดงความเที่ยงตรงของแบบทดสอบ (Validity)

หมายถึง หลักฐานที่แสดงว่าแบบทดสอบวัดในสิ่งที่ต้องการวัดได้อย่างถูกต้องตรงตามความมุ่งหมายของเนื้อหาวิชาวิทยาศาสตร์ ในการศึกษาครั้งนี้เป็นการแสดงหลักฐานความเที่ยงตรงเชิงเนื้อหา (Content Validity) ซึ่งผู้วิจัยสร้างให้ครอบคลุมเนื้อหาตามตารางวิเคราะห์รายละเอียดและประมาณค่าจากวิธีของโรวินเนลลี และแฮมเบิลตัน (บุญเชิด ภิญญอนันตพงษ์. 2545 : 95 ; อ้างอิงจาก Rovineli Hambleton. 1977) คือให้ผู้เชี่ยวชาญ จำนวน 5 ท่านเป็นผู้ตรวจสอบ

10. การเตา

หมายถึง คะแนนข้อที่นักเรียนตอบถูกหลังข้อที่นักเรียนตอบผิดในแต่ละตำแหน่งของการตอบผิดตำแหน่งต่าง ๆ

11. ผู้เชี่ยวชาญ

หมายถึง ผู้ที่มีคุณวุฒิตั้งแต่ปริญญาโททางการวัดผลการศึกษาไม่น้อยกว่า 5 ปี จำนวน 3 ท่าน และ ผู้ที่มีประสบการณ์การสอนทางด้านวิชาวิทยาศาสตร์เป็นเวลาไม่น้อยกว่า 5 ปี จำนวน 2 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาค้นคว้าครั้งนี้จะกล่าวถึงเอกสารและงานวิจัยดังต่อไปนี้ คือ

- ✓ 1. วิธีการตรวจให้คะแนน
- 2. การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้
- ✓ 3. การเรียงข้อคำถามในแบบทดสอบ
- 4. ความเที่ยงตรงของแบบทดสอบ
- 5. ความเชื่อมั่นของแบบทดสอบ
- 6. เอกสารที่เกี่ยวข้องกับการสอนวิทยาศาสตร์ในโรงเรียนประถมศึกษา
- 7. งานวิจัยที่เกี่ยวข้องกับวิธีการตรวจให้คะแนน
- 8. งานวิจัยที่เกี่ยวข้องกับการเรียงข้อคำถามในแบบทดสอบ
- 9. สมมติฐานในการศึกษาค้นคว้า

1. วิธีการตรวจให้คะแนน

ในปัจจุบันมีผู้ให้แนวคิดเกี่ยวกับวิธีการตรวจให้คะแนนแบบทดสอบชนิดต่าง ๆ ที่ใช้ในกระบวนการวัดผลการศึกษา เพื่อสามารถบอกความหมายของคะแนนที่ได้ตรงตามสภาพจริงของนักเรียน หรือใกล้เคียงความสามารถที่แท้จริงให้มากที่สุดไว้หลายวิธี ดังนี้

1. วิธีการตรวจให้คะแนนของคัมบัส (Coombs. 1956 : 308 - 313)

คัมบัสเสนอให้มีวิธีการตอบและตรวจให้คะแนน โดยให้เลือกคำตอบที่ผิดแทนการเลือกคำตอบที่ถูกต้องพร้อมให้ความเห็นว่าการเลือกคำตอบที่ผิดซึ่งมีอยู่หลายตัวเลือกในคำถามหนึ่ง ๆ นั้น สำหรับการตรวจให้คะแนนนั้น จะให้เป็นรายตัวเลือก ถ้าตัวเลือกที่ผิดจริงจะได้คะแนนตัวเลือกละ 1 คะแนน ถ้าเลือกตัวเลือกที่ถูกต้องจะได้ 1 - K คะแนน (เมื่อ K คือจำนวนตัวเลือกทั้งหมดใน 1 ข้อ)

2. วิธีการตรวจให้คะแนนของเดรสเซล และ ชมิท (Dressel and Schmid. 1953 : 574 - 595)

เดรสเซล และชมิทเสนอวิธีให้คะแนนความมั่นใจ วิธีนี้จะเปิดโอกาสให้ผู้ตอบเลือกตอบอย่างเสรี โดยที่จะเลือกตอบกี่ตัวเลือกก็ได้ และจะต้องบอกถึงระดับความมั่นใจของตัวเลือกที่เลือกถูกต้องเพียงไร วิธีบอกความมั่นใจของเดรสเซล และชมิทมีการแบ่งระดับความมั่นใจละเอียดเป็น 5 ระดับโดยใช้กับแบบทดสอบเลือกตอบ 4 แบบ คือ 1) แบบเลือกตอบเสรี (Free Choice Test) เป็นแบบเลือกตอบที่ผู้สอบจะเลือกคำตอบที่ข้อก็ได้ ตามที่ตนคิดว่าถูกต้อง 2) แบบสอบบอกความมั่นใจ (Degree of Certainly Test) เป็นแบบทดสอบที่ผู้สอบเลือกตอบได้หนึ่งข้อและบอกว่าตนเองแน่ใจในคำตอบนั้นอยู่ในระดับใดโดยแบ่งความมั่นใจออกเป็น 4 ระดับ 3) แบบมีคำตอบถูกมากกว่าหนึ่ง (Multiple Answer Test) เป็นแบบทดสอบที่มีจำนวนคำตอบถูกในแต่ละข้อไม่จำกัด ผู้สอบจะต้องทำเครื่องหมายทุกข้อที่ตนคิดว่าถูก 4) แบบมีคำตอบถูกมากกว่า 2 คำตอบ (Two Answer Test) จาก 5 ตัวเลือกในแต่ละข้อ ดังนี้

ตาราง 1 วิธีการให้คะแนนแบบเลือกเสรี

จำนวนคำตอบที่ทำ เครื่องหมาย	คะแนน	
	ผู้ทำเครื่องหมายข้อถูก	ผู้ไม่ได้ทำเครื่องหมายข้อถูก
1	4	-1
2	3	-2
3	2	-3
4	1	-4
5	0	-5

ตาราง 2 วิธีให้คะแนนแบบบอกระดับความมั่นใจ

ระดับความมั่นใจ	ตอบถูกได้คะแนน	ตอบผิดได้คะแนน
แน่ใจ	4	-4
ค่อนข้างแน่ใจ	3	-3
เดาอย่างมีเหตุผล	2	-2
ไม่แน่ใจ	1	-1

ข้อจำกัดของวิธีนี้ คือ กรณีที่ผู้ตอบเลือกทุกตัวเลือกจะต้องมีข้อถูกอยู่ด้วย 1 ข้อ ในกรณีนี้คะแนนลบห้า จึงไม่น่าจะเกิดขึ้น และการให้คะแนนติดลบไม่สอดคล้องกับการวัดทางจิตวิทยา

3. วิธีการตรวจให้คะแนนของเดวิส และ ฟิฟเฟอร์ (พิมพร ภูมิรา. 2544 : 13 อ้างอิงจาก Davis and Fiffer. 1959 : 159 - 169)

เดวิส และ ฟิฟเฟอร์ เสนอให้มีการตรวจให้คะแนนแบบให้น้ำหนักแก่ทุกตัวเลือกไม่เท่ากัน โดยให้ผู้เชี่ยวชาญหลาย ๆ ท่านเป็นผู้กำหนดน้ำหนัก ซึ่งเป็นวิธีหนึ่งที่จะช่วยให้คะแนนจากแบบทดสอบมีความเชื่อมั่นที่สูงขึ้น แต่ยุ่งยากในทางปฏิบัติจึงไม่เหมาะสมที่จะนำมาใช้ในแบบทดสอบที่ครูสร้างขึ้น

4. วิธีตรวจให้คะแนนของเฟรดเดอริก บี เดวิส (Davis. 1966 : 79) ได้เสนอสูตรการตรวจให้คะแนนเพื่อแก้การเดาของผู้สอบ

$$X_c = R - \frac{W}{C - 1}$$

เมื่อ	X_c	แทน	คะแนนที่แก้เดา
	R	แทน	จำนวนข้อที่ตอบถูก
	W	แทน	จำนวนข้อที่ตอบผิด
	C	แทน	จำนวนตัวเลือกของข้อสอบแต่ละข้อ

จากสูตรดังกล่าวเป็นวิธีแก้การเดา โดยการหักสัดส่วนของจำนวนข้อที่ตอบผิดออกจากจำนวนข้อที่ตอบถูก วิธีนี้เคยถูกวิจารณ์ว่ามีรากฐานมาจากข้อตกลงเบื้องต้นที่ยังเป็นปัญหา คือคำตอบผิดทุกข้อถือว่าเป็นผลเนื่องมาจากการเดา จากข้อเท็จจริงที่ว่าข้อสอบที่ตอบผิดนั้น บางครั้งผู้สอบไม่ได้เดา แต่เป็นเพราะได้รับความรู้ที่ผิด หรือข้อสอบนั้นมีตัวลวงที่มีเหตุผลดี (เพ็ญศรี สว่างเนตร. 2520 : 2)

5. วิธีการตรวจให้คะแนนของอีเบล (Ebel. 1965 : 224)

อีเบล เสนอวิธีการตรวจให้คะแนนวิธีนี้เพื่อแก้การเดา คือ แทนที่จะหักคะแนนข้อที่ตอบผิดกลับเพิ่มคะแนนเท่ากับข้อที่ไม่ตอบโดยถือว่าข้อที่เว้นไว้นั้นเป็นข้อที่ผู้ตอบไม่รู้ หรือยังทำไม่ทันแต่ถ้าเขาตอบโดยการเดา หรือทำเครื่องหมายลงในกระดาษคำตอบโดยไม่ได้อ่านข้อสอบแล้วมีโอกาสที่จะตอบถูกหนึ่งในจำนวนตัวเลือกของข้อนั้น ใช้สูตรแก้การเดาเป็น

$$X_c = R + \frac{O}{C}$$

เมื่อ	X_c	แทน	คะแนนที่แก้การเดา
	R	แทน	จำนวนข้อที่ตอบถูก
	O	แทน	จำนวนข้อที่เว้น
	C	แทน	จำนวนตัวเลือกของข้อสอบแต่ละข้อ

6. วิธีตรวจให้คะแนนแบบกิบบอนส์ โอล์คิน และโซเบล (Gibbons, Olkin and Sobel. 1979 : 261)

กิบบอนส์ และ คณะ ได้เสนอวิธีการตอบ และการตรวจให้คะแนนในแต่ละข้อ โดยให้ผู้สอบเลือกตอบตัวเลือกที่ถูกได้ตั้งแต่ 1 ตัวเลือกจนถึง $k - 1$ ตัวเลือก เมื่อ k คือจำนวนตัวเลือกทั้งหมดในแต่ละข้อ ส่วนวิธีการตรวจให้คะแนนยึดถือหลักว่าคะแนนคาดหวัง (Expected Score) ที่ได้จากการเดาสุ่มเป็นศูนย์ ในกรณีที่ข้อสอบมี 5 ตัวเลือกและผู้สอบตอบอย่างสุ่ม 2 ตัวเลือก ความน่าจะเป็นที่เขาจะตอบถูกเป็น $2 / 5$ และความน่าจะเป็นที่เขาจะตอบผิดเป็น $3 / 5$ คะแนนคาดหวังจะเป็น $E(S) = 3(2 / 5) - 2(3 / 5)$ ในกรณีนี้ถ้าเขาตอบถูกก็จะได้คะแนน 3 คะแนน แต่ถ้าตอบผิดจะได้คะแนน -2 คะแนน ซึ่งรูปแบบการให้คะแนนสามารถแสดงได้ ดังนี้

ตาราง 3 วิธีให้คะแนนแบบกิบบอนส์ โอล์คิน และโซเบล

จำนวนตัวเลือกที่ตอบ	คะแนนเมื่อตอบรวมเอาตัวเลือกที่ถูก	คะแนนเมื่อตอบไม่รวมเอาตัวเลือกที่ถูก
0	-	0
1	4	-1
2	3	-2
3	2	-3
4	1	-4
5	0	

จะเห็นว่าการตอบและวิธีการให้คะแนนตามวิธีของกิบบอนส์ และคณะสามารถวัดคนที่มีความรู้
น้อยๆ ไปจนถึงคนที่มีความรู้จริงในเรื่องนั้น ๆ ทั้งยังสามารถแก้การเดาสุ่มได้อีกด้วย

7. วิธีการตรวจให้คะแนนของเกริกชัย ฮวบเจริญ (2525 : 5 - 6)

วิธีนี้ผู้สอบสามารถรู้ได้ทันทีว่าคำตอบที่ตนตอบ ถูก หรือ ผิด เพราะตอบโดยใช้เครื่องบันทึก
คำตอบและปากกาถูกลิ้นจิ้มลงไปในช่องวงกลมที่มีอักษร ก ข ค ง และ จ คือถ้าผู้ตอบเลือกคำตอบที่ถูก
ต้องปากกาก็จะสามารถจิ้มลึกลงไปถึงแผ่นล่างได้ ถ้าจิ้มไม่ถึงแผ่นล่างแสดงว่าเป็นคำตอบที่ไม่ถูกต้อง ผู้สอบ
จะทำข้อเดิมต่ออีกหรือไม่ก็ได้ ถ้าจะทำข้อเดิมต่อก็ให้กลับไปอ่านคำถามแล้วเลือกจิ้มในช่องวงกลมที่เหลือถ้า
ยังทำไม่ถูกอีกอาจตัดสินใจเลือกหรือไม่เลือกอีกก็ได้ ทำเช่นนี้เรื่อยไป การตรวจให้คะแนนใช้วิธี ดังนี้

ตาราง 4 วิธีการตรวจให้คะแนนของเกริกชัย ฮวบเจริญ

จำนวนครั้งที่เลือก	คะแนนที่ได้เมื่อตอบถูก	คะแนนที่ได้เมื่อตอบผิด
1	4	-1
2	2	-2
3	0	-3
4	-2	-4
5	-4	

8. วิธีการตรวจให้คะแนนของสำราญ มีแจ้ง (2525 : 10 - 11)

สำราญ มีแจ้ง เสนอวิธีตอบและตรวจให้คะแนน โดยให้ผู้ตอบเลือกตอบตัวเลือกที่ถูกที่สุดเพียง
ตัวเลือกเดียวสำหรับข้อที่มั่นใจว่าถูกแน่นอน ส่วนข้อที่ผู้ตอบไม่มั่นใจก็ให้ตอบได้ 2 หรือ 3 ตัวเลือก ที่คิดว่า
ถูก สำหรับการตรวจให้คะแนนจะให้คะแนนลดลงตามจำนวนตัวเลือกที่ตอบ ถ้าตอบเพียงตัวเลือกเดียวแล้ว
ถูกจะได้ 4 คะแนน ถ้าผิดจะได้ 0 คะแนน ถ้าตอบ 2 หรือ 3 ตัวเลือกแล้วถูก (มีตัวเลือกถูกอยู่ในชุดที่เลือก)
ก็จะได้คะแนน 2 และ 1 คะแนนตามลำดับ ถ้าผิดจะได้ 0 คะแนน ดังนี้

ตาราง 5 วิธีการให้คะแนนของสำราญ มีแจ้ง

จำนวนตัวเลือกที่ตอบ คะแนนเมื่อตอบรวม	เอาตัวเลือกที่ถูก	คะแนนเมื่อตอบไม่รวม เอาตัวเลือกที่ถูก
0	-	0
1	4	0
2	2	0
3	1	0
4	0	0
5	0	0

จะเห็นว่า วิธีการให้คะแนนของ สำราญ มีแจ้ง ไม่สามารถวัดคนที่มีความรู้น้อยได้จริงเพราะ คนที่สามารถกำจัดตัวลวงได้ 1 ตัวก็แสดงว่าเขามีความรู้ส่วนหนึ่งแล้วนอกจากนี้ยังทำให้ผู้ที่ยังคลุมเครือ ระหว่าง 3 และ 2 ตัวเลือกว่าตัวเลือกใดถูกเสี่ยงตอบเพียงตัวเลือกเดียว เพราะถ้าตอบถูกเขาจะได้คะแนน เพิ่มขึ้นจากการเสี่ยงเป็น 4 เท่า และ 2 เท่าของคะแนนที่เขาควรจะได้ตามลำดับทั้งยังเปิดโอกาสให้ผู้ตอบที่ เดาสุ่มได้คะแนนอีกด้วย

9. วิธีการตรวจให้คะแนนของอนันต์ ศรีโสภา (2526 : 13 - 19)

อนันต์ ศรีโสภา เสนอวิธีการตอบวิธีใหม่โดยให้ตอบทุกตัวเลือกการตอบวิธีนี้ให้ผู้สอบระบุทั้ง คำตอบที่ถูก และคำตอบที่ผิด วิธีการตรวจให้คะแนนจะให้ 1 คะแนน ถ้าตัวเลือกที่ถูกจริงหรือผิดจริงแต่ ถ้าเลือกตอบตรงข้ามกับเฉลยจะได้ตัวเลือกละ -1 คะแนน เช่น ให้ทำเครื่องหมาย + ลงบนตัวเลือกที่เห็นว่า ถูก และทำเครื่องหมาย - ลงบนตัวเลือกที่เห็นว่าผิด ถ้าตัวเลือกใดที่ไม่รู้ว่าถูก หรือผิดให้ข้ามไปไม่ต้องทำ เครื่องหมายใด ๆ การตรวจให้คะแนน ถ้าผู้สอบทำเครื่องหมาย + ลงบนตัวเลือกที่ถูก หรือเครื่องหมาย - ลงบนตัวเลือกที่ผิดจะได้ตัวเลือกละ 1 คะแนน แต่ถ้าทำเครื่องหมาย + ลงบนตัวเลือกที่ผิด หรือเครื่องหมาย - ลงบนตัวเลือกที่ถูก จะได้ตัวเลือกละ -1 คะแนน ส่วนตัวเลือกที่ผู้สอบไม่ได้ทำเครื่องหมายใด ๆ เลยจะได้ 0 คะแนน ซึ่งการตรวจให้คะแนนวิธีนี้เป็น การวัดความรู้ในข้อคำถามนั้น ๆ ได้สมบูรณ์ยิ่งกว่าการ เลือกคำตอบที่ถูกเพียงคำตอบเดียว และยังสามารถทราบความรู้ที่ผิดของผู้สอบได้ดีกว่าวิธีตอบแบบ ธรรมดาอีกด้วย

10. วิธีการตรวจให้คะแนนแบบธรรมดา (The Conventional Method)

เป็นวิธีการตรวจให้คะแนนที่นิยมปฏิบัติกันแบบทดสอบเลือกตอบ ที่หมายถึง วิธีการให้ผู้สอบ เลือกตอบเฉพาะตัวเลือกที่ถูกให้ 1 คะแนน และให้ 0 คะแนน ถ้าผู้ตอบเลือกตัวเลือกที่ผิดหรือไม่เลือก ตอบเลย บุญชม ศรีสะอาด (2540 : 122) กล่าวถึงการตรวจให้คะแนนตามวิธีธรรมดาว่าคือการให้ 1 คะแนน สำหรับคำตอบถูก และ 0 คะแนนสำหรับคำตอบผิด วิธีนี้จึงเรียกอีกอย่างหนึ่งว่า "วิธี 0 - 1"

หรือ “วิธีให้ข้อถูกต้อง 1 คะแนน” หรือ “วิธีให้คะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิม” แต่เนื่องจากเป็นวิธีที่ใช้กันโดยทั่วไปในโรงเรียนและสถาบันการศึกษาจนถือเป็นปกติธรรมดาจึงเรียกว่าเป็น “วิธีธรรมดา”

11. วิธีการตรวจให้คะแนนโดยกฎชิลลิ่ง (Ceiling Rules)

จากการศึกษางานวิจัยในต่างประเทศและในประเทศ มีผู้ศึกษาวิจัยเกี่ยวกับผลการใช้กฎชิลลิ่งเพื่อหาความสามารถที่แท้จริงของผู้สอบ ดังนี้

สมบุรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) ได้ทำการศึกษาวิจัยเปรียบเทียบผลการใช้กฎชิลลิ่งจากการกำหนดตำแหน่งการตอบผิดที่แตกต่างกันต่อค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์โดยใช้ความยากของข้อสอบเป็นเกณฑ์ในการพิจารณาการเดาตอบถูกของผู้สอบ ดังนั้นเมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุดตามรูปแบบของกฏเดาแล้วนักเรียนที่ตอบข้อสอบข้อที่ยากได้ถูกต้องในขณะที่ตอบข้อสอบข้อที่ง่ายผิด ความน่าจะเป็นในการตอบถูกนั้นควรเป็นไปโดยการเดามากกว่าตอบถูกด้วยความสามารถที่แท้จริง จึงไม่ควรได้คะแนนในข้อที่ตอบถูกนั้น แต่การตอบถูกในข้อต่อไปซึ่งเป็นข้อที่ยากกว่า จะสรุปโดยทันทีไม่ได้ว่านักเรียนทุกคนตอบถูกโดยการเดา เพราะนักเรียนบางคนที่ตอบข้อสอบข้อที่ง่ายผิดแต่ตอบข้อสอบข้อที่ยากกว่าถูกนั้นอาจมีสาเหตุมาจากความบกพร่องในการคิด จึงทำให้การประมาณค่าความสามารถที่แท้จริงไม่ถูกต้องนัก และเป็นการไม่ยุติธรรมสำหรับนักเรียน สมบุรณ์ สุริยวงศ์ จึงเสนอวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือ การตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้า กับการไม่กำหนดตำแหน่งการตอบผิดมากำหนดความสามารถที่แท้จริงของผู้สอบ

วิลสัน และคณะ (Wilson , et al. 1988 : 213 - 216) เสนอใช้ระดับความยากของข้อสอบเป็นเกณฑ์ในการพิจารณาการเดาตอบถูกของผู้สอบ เพราะผู้สอบแต่ละคนจะมีความสามารถระดับหนึ่ง ที่จะทำข้อสอบข้อที่มีความยากระดับต่าง ๆ ได้ถูกต้องตามความสามารถของแต่ละคน ดังนั้นเมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุดแล้ว นักเรียนที่ตอบข้อสอบข้อที่ยากได้ถูกต้องในขณะที่ตอบข้อสอบข้อที่ง่ายผิดความน่าจะเป็นในการตอบถูกนั้นควรเป็นไปโดยการเดามากกว่าตอบถูกด้วยความสามารถที่แท้จริง นักเรียนจึงไม่ควรได้คะแนนในข้อที่ตอบถูกนั้นแต่การตอบถูกในข้อต่อไปซึ่งเป็นข้อที่ยากกว่าจะสรุปโดยทันทีไม่ได้ว่านักเรียนทุกคนตอบถูกโดยการเดา เพราะนักเรียนบางคนที่ตอบข้อสอบข้อที่ง่ายผิดแต่ตอบข้อสอบข้อที่ยากกว่าถูกอาจมีสาเหตุมาจากความบกพร่องในการคิด เช่นนักเรียนบางคนคิดมากในข้อสอบบางข้อหรืออาจเกิดจากความสะเพร่าทำให้ตอบผิดทั้ง ๆ ที่มีความรู้ จึงทำให้การประมาณค่าความสามารถที่แท้จริงไม่ถูกต้องนัก และเป็นการไม่ยุติธรรมสำหรับนักเรียน วิลสัน และคณะ จึงเสนอการตรวจให้คะแนนโดยใช้กฎชิลลิ่งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือ การตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้า กับการไม่กำหนดตำแหน่งการตอบผิดมากำหนดความสามารถที่แท้จริงของผู้สอบ

จินตนา ต่อทรัพย์สินชัย (2533) อัมพร ยิ้มแย้ม (2533) และมานะ คิ้วเจริญ (2539) ศึกษาวิธีการตรวจให้คะแนนโดยกฎชิลลิ่ง ของ สมบุรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) วิลสัน และคณะ (Wilson , et al. 1988 : 213 - 216) แล้วจึงทำการศึกษาวิจัยกับนักเรียนชั้นมัธยมศึกษา เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่เรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด โดยกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้า กับการไม่กำหนดตำแหน่งการตอบผิด มากำหนดความสามารถที่แท้จริงของผู้สอบ มีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

1. การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อแรก (0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน
2. การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สอง (0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน
3. การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สาม (0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน
4. การตอบผิดตำแหน่งที่สี่ หมายถึงการกำหนดให้คะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สี่ (0,0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน
5. การตอบผิดตำแหน่งที่ห้า หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่ห้า (0,0,0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน
6. การไม่กำหนดตำแหน่งการตอบผิด หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อสุดท้าย

2. การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้

การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้นั้นมีความสำคัญมากทำให้สามารถวัดผลการเรียนรู้ได้ชัดเจน ในที่นี้จะเสนอแนวคิดในการเขียนข้อสอบวัดพฤติกรรมการเรียนรู้ด้านพุทธิพิสัยที่สำคัญๆ ดังนี้

บลูม และคณะ (ลัวัน สายยศและ อังคณา สายยศ 2539 : 41 - 44 ; อ้างอิงจาก Bloom , et al. 1956.) ได้แบ่งพฤติกรรมการเรียนรู้ออกเป็น 6 พฤติกรรมย่อย ๆ ดังนี้

1. ความรู้ (Knowledge) เป็นความสามารถในการระลึกในสิ่งที่ได้เรียนรู้มาแล้ว หรือก็คือความจำ ซึ่งแบ่งเป็นข้อย่อย ๆ ได้ ดังนี้

1.1 ความรู้ด้านเนื้อหา (Knowledge of Specifics) เป็นความสามารถในการจดจำเนื้อหาของสิ่งที่เรียนหรือประสบพบมา แบ่งเป็น 2 อย่าง คือ

1.1.1 ความรู้เกี่ยวกับศัพท์และนิยาม (Knowledge of Terminology) ความรู้แบบนี้เป็นสัญลักษณ์ ศัพท์ นิยามที่ตกลงกันไว้เพื่อใช้ในเรื่องใดเรื่องหนึ่ง เพื่อให้เป็นความหมายที่สะดวก

1.1.2 ความรู้เกี่ยวกับความจริง (Knowledge of Specific Facts) เป็นความสามารถในการจดจำสิ่งที่เป็นความจริงที่เรียนรู้มา ความจริงในที่นี้เป็นลักษณะ วันที่ เดือน ปี สถานที่ บุคคล เหตุการณ์ ฯลฯ ที่เกิดขึ้นเป็นจริงมาแล้ว

1.2 ความรู้เกี่ยวกับวิธีดำเนินการในเนื้อหา (Knowledge of Ways and Means of Dealing with Specifics) ความสามารถด้านนี้ เป็นความจำในด้านวิธีการจัดระบบจัดการศึกษาพิจารณา วิพากษ์ วิวิจารณ์ รวมทั้งวิธีแสวงหาความรู้และลำดับขั้นของเวลา แบ่งเป็น 5 อย่าง คือ

1.2.1 ความรู้เกี่ยวกับระเบียบประเพณี (Knowledge of Conventions) เป็นความสามารถในการจดจำประเพณี วัฒนธรรม ธรรมเนียมหรือการกระทำที่เป็นนิสัยยึดถือกันในสังคม หรือในเนื้อหาวิชานั้น ๆ

1.2.2 ความรู้เกี่ยวกับแนวโน้มและลำดับขั้น (Knowledge of Trends and Sequences) เป็นความสามารถในการจำ เพื่อหาส่วนที่เกี่ยวข้องกับการแสดงออกทางแนวโน้มและลำดับขั้นตอนในเรื่องใดเรื่องหนึ่ง

1.2.3 ความรู้เกี่ยวกับการจำแนกประเภท (Knowledge of Classifications and Categories) เป็นความจำในเรื่องจัดประเภท กลุ่มชุดของความรู้ในเนื้อหาวิชาที่เรียนรู้อยู่แล้ว

1.2.4 ความรู้เกี่ยวกับเกณฑ์ (Knowledge of Criteria) หมายถึง ความจำเกณฑ์ต่าง ๆ ในการเกิดหลักการ มโนภาพ ความคิดเห็น และอื่น ๆ

1.2.5 ความรู้เกี่ยวกับวิธีการ (Knowledge of Methodology) เป็นลักษณะการจำวิธีการในการค้นหาความรู้ จำเทคนิค และกระบวนการต่าง ๆ ที่เคยเรียนมาแล้ว

1.3 ความรู้เกี่ยวกับความคิดรวบยอด (Knowledge of the Universals and Abstractions in a Field) ความจำแบบนี้เป็นความจำขั้นสูงขึ้นไป แบ่งออกเป็น 2 อย่าง คือ

1.3.1 ความรู้เกี่ยวกับหลักวิชาและขยายนัยทั่วไป (Knowledge of Principles and Generalizations) เมื่อสอนหลักวิชาและขยายนัยทั่วไปในหลักวิชานั้น ๆ แล้ว จุดประสงค์นี้ต้องการจะให้จำสิ่งนั้น ๆ ให้ได้

1.3.2 ความรู้เกี่ยวกับทฤษฎี และโครงสร้าง (Knowledge of Theories and Structures) ระดับนี้จุดประสงค์ เพื่อให้ผู้เรียนสามารถจำทฤษฎี และ โครงสร้างของสิ่งที่เรียนมาแล้วให้ได้

2. ความเข้าใจ(Comprehension) หมายถึง ความสามารถในการแปลความ ตีความและขยายความจากสื่อความหมายต่าง ๆ ที่ได้พบเห็น แบ่งส่วนย่อยออก 3 อย่าง คือ

2.1 การแปลความ (Translation) เป็นความสามารถในการถ่ายทอดความหมายจากภาษาระดับหนึ่งมาเป็นอีกระดับหนึ่งให้เข้าใจกันง่ายขึ้น

2.2 การตีความ (Interpretation) เป็นความสามารถในการสรุป การแปลความ มองภาพส่วนรวมมาเป็นใจความสั้น ๆ อย่างได้ความ

2.3 การขยายความ (Extrapolation) เป็นความสามารถในการทำนาย หรือคาดคะเนข้อเท็จจริงล่วงหน้าโดยอาศัยแนวโน้มที่มีมาแล้ว

3. การนำไปใช้ (Application) เป็นความสามารถในการนำหลักวิชาไปใช้แก้ปัญหาในสถานการณ์ใหม่

4. การวิเคราะห์ (Analysis) เป็นความสามารถในการแยกแยะ เพื่อหาส่วนย่อย ๆ ของเหตุการณ์เรื่องราวหรือเนื้อหาต่าง ๆ ว่าประกอบด้วยอะไร มีความสามารถอย่างไร อะไรเป็นเหตุ อะไรเป็นผล และที่เป็นไปอย่างนั้นอาศัยหลักการอะไร การวิเคราะห์แบ่งย่อยเป็น 3 อย่าง คือ

4.1 วิเคราะห์ความสำคัญ (Analysis of Elements) เป็นการวิเคราะห์ว่าสิ่งที่อยู่กันนั้นอะไรสำคัญ หรือจำเป็น หรือมีบทบาทที่สุด ตัวไหนเป็นเหตุ ตัวไหนเป็นผล

4.2 วิเคราะห์ความสัมพันธ์ (Analysis of Relationships) เป็นการหาความสัมพันธ์หรือความเกี่ยวข้องของส่วนย่อยในปรากฏการณ์ หรือเนื้อหาหนึ่งเพื่อนำมาอุปมาอุปไมย

4.3 วิเคราะห์หลักการ (Analysis of Organizational Principles) เป็นความสามารถที่จะจับเค้าเงื่อนของเรื่องราวเห็นว่ายึดหลักการใด มีเทคนิค หรือยึดคติใด

5. การสังเคราะห์ (Synthesis) เป็นความสามารถในการผสมผสานส่วนย่อยต่าง ๆ เข้าด้วยกัน เพื่อให้เป็นสิ่งใหม่อีกรูปแบบหนึ่งมีคุณลักษณะโครงสร้างหรือหน้าที่ใหม่แปลกแตกต่างไปจากของเดิมแบ่งย่อยออกเป็น 3 อย่าง คือ

5.1 สังเคราะห์ข้อความ (Production of Unique Communication) เป็นความสามารถในการสังเคราะห์ข้อความโดยสื่อหรือโดยการพูด การเขียน การวิพากษ์วิจารณ์ หรือหาข้อยุติบางประการ

5.2 สังเคราะห์แผนงาน (Production of Plans and Proposed Set of Operations) เป็นความสามารถด้านการวางแผนหรือการจัดกระทำการออกแบบหรือการกระทำในลักษณะเดียวกันนี้

5.3 สังเคราะห์ความสัมพันธ์ (Derivation of a Set of Abstract Relation) เป็นความสามารถในการนำนามธรรมย่อย ๆ มาสัมพันธ์กัน เกิดเป็นทฤษฎี สมมติฐาน สูตร หรือ กฎขึ้น

6. การประเมินค่า (Evaluation) เป็นความสามารถในการพิจารณาตัดสินลงสรุป เกี่ยวกับคุณค่าของความคิดทุกชนิดเพื่อเปรียบเทียบเกณฑ์หรือมาตรฐานที่กำหนดให้ แบ่งย่อยออกเป็น 2 อย่าง คือ

6.1 ประเมินค่าโดยอาศัยข้อเท็จจริงภายใน (Judgement in terms of Internal Evidence) การประเมินแบบนี้พิจารณาหาความถูกต้อง สมเหตุสมผล ความสอดคล้องโดยอาศัยเกณฑ์ภายในของสิ่งนั้น เป็นสำคัญ

6.2 ประเมินค่าโดยอาศัยเกณฑ์ภายนอก (Judgement in terms of External Criteria) การประเมินแบบนี้อาศัยเกณฑ์หรือมาตรฐานจากภายนอกเอาไว้มเปรียบเทียบ เกณฑ์เหล่านี้อาจเป็นเกณฑ์ที่สังคม หรือระเบียบประเพณีกำหนดไว้ก็ได้

นอกจากนี้ โกรว์ (Gorow. 1966 : 41 - 43) ยังได้อธิบายถึงความสามารถของแบบทดสอบวัดผลสัมฤทธิ์ที่สามารถใช้วัดพฤติกรรมการเรียนรู้ในชั้นต่างๆต่อไปนี้

1. การเรียนรู้คำศัพท์ (Knowledge of Terminology)
2. การเรียนรู้ข้อเท็จจริง (Knowledge of Specific Facts)
3. การเรียนรู้แบบแผน (Knowledge of Conventions)
4. การเรียนรู้แนวโน้มของการเกิดเหตุการณ์ (Knowledge of Trends and Sequences)
5. การเรียนรู้จำพวกและประเภท (Knowledge of Classification and Categories)
6. การเรียนรู้เกณฑ์ในการตัดสินใจ (Knowledge of Criteria for Making Judgement)
7. การเรียนรู้วิธีการ (Knowledge of Method)

คลอปเฟอร์ (ภพ เลหาไพบูลย์. 2542 : 99 – 110 ; อ้างอิงจาก Leopold E. Klopfer. n.d.)

เป็นนักการศึกษาสาขาวิทยาศาสตร์ในมหาวิทยาลัยพิตสเบิร์ก สหรัฐอเมริกา คลอปเฟอร์ได้ศึกษาวัตถุประสงค์ทางการศึกษาของบลูม แล้วนำมากำหนดเป็นวัตถุประสงค์ให้เหมาะสมกับการเรียนการสอนวิทยาศาสตร์เพื่อให้ได้ทั้งเนื้อหาที่เป็นความรู้ทางวิทยาศาสตร์และพฤติกรรมที่ต้องการให้เกิดขึ้นในตัวผู้เรียน

วัตถุประสงค์การเรียนการสอนวิทยาศาสตร์ตามแนวลอปปเฟอร์เป็นที่แพร่หลายมากในปัจจุบัน วัตถุประสงค์การเรียนการสอนวิทยาศาสตร์ตามแนวลอปปเฟอร์มีดังนี้

1. ความรู้ความเข้าใจ (Knowledge and Comprehension) ความรู้ความเข้าใจวิทยาศาสตร์นั้น ผู้เรียนอาจได้รับมาจากกระบวนการค้นคว้าทางวิทยาศาสตร์ แบ่งได้เป็น ความรู้วิทยาศาสตร์ และความเข้าใจวิทยาศาสตร์

1.1 ความรู้วิทยาศาสตร์ หมายถึง เนื้อหาที่เป็นความรู้ทางวิทยาศาสตร์ ซึ่งแบ่งเป็น

9 ประการ คือ

1.1.1 ความรู้เกี่ยวกับข้อเท็จจริง

1.1.2 ความรู้เกี่ยวกับคำศัพท์ทางวิทยาศาสตร์

1.1.3 ความรู้เกี่ยวกับมโนคติทางวิทยาศาสตร์

1.1.4 ความรู้เกี่ยวกับข้อตกลง

1.1.5 ความรู้เกี่ยวกับแนวโน้มและลำดับขั้นตอน

1.1.6 ความรู้เกี่ยวกับการจำแนกประเภท จัดประเภทและเกณฑ์

1.1.7 ความรู้เกี่ยวกับเทคนิคและกรรมวิธีทางวิทยาศาสตร์

1.1.8 ความรู้เกี่ยวกับหลักการและกฎวิทยาศาสตร์

1.1.9 ความรู้เกี่ยวกับทฤษฎีทางวิทยาศาสตร์

1.2 ความเข้าใจวิทยาศาสตร์ เป็นการใช้ความคิดที่สูงกว่าความจำ แบ่งเป็น 2 ประเภท คือ

1.2.1 การนำความรู้ไปใช้ในสิ่งใหม่

1.2.2 การแปลความหมายของความรู้ในรูปของสัญลักษณ์หนึ่งไปเป็นรูปของ

อีกสัญลักษณ์หนึ่ง

2. กระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ (Process of Scientific Inquiry) ผู้เรียน ได้แสดงพฤติกรรมถึงการมีส่วนร่วมในการสืบเสาะหาความรู้ด้วยตนเอง เป็นกระบวนการที่นักวิทยาศาสตร์ใช้สำหรับการศึกษาเรื่องราวของธรรมชาติและสร้างสรรค์แนวความคิดใหม่ๆ ขึ้นมา กระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์มีดังนี้ คือ

2.1 การสังเกตและการวัด การสังเกตเป็นการใช้ประสาทสัมผัสทั้งห้าเข้าไปสำรวจวัตถุหรือปรากฏการณ์ธรรมชาติโดยตรง ถ้าใช้การสังเกตอย่างเดียวไม่สามารถบอกปริมาณที่ถูกต้องแน่นอนได้ ต้องใช้ทั้งการสังเกตและการวัดควบคู่กันไป

2.2 การมองเห็นปัญหาและหาทางที่จะแก้ปัญหา ประกอบด้วยความสามารถย่อยๆ คือ การยอมรับและมองเห็นปัญหา การตั้งสมมติฐาน การเลือกวิธีการที่เหมาะสมในการทดสอบสมมติฐาน และการออกแบบการทดลองที่เหมาะสมสำหรับการทดสอบสมมติฐาน

2.3 การแปลความหมายข้อมูลและการสร้างข้อสรุป ข้อมูลที่ผู้เรียนได้จากการทดลองนั้น เป็นการบันทึกผลของการสังเกตและการวัดต่างๆซึ่งข้อมูลเหล่านั้นจะต้องถูกจัดกระทำต่อไปเพื่อให้ได้ผลลัพธ์ที่มีคุณค่าสูงขึ้นในการศึกษาเรื่องนั้น

2.4 การสร้าง การทดสอบ และการปรับปรุงแบบจำลองเชิงทฤษฎี ซึ่งแบบจำลองเชิงทฤษฎีสามารถอธิบายปรากฏการณ์ หลักการหรือกฎเกณฑ์ต่างๆ ที่ค้นพบ และถ้าแบบจำลองเชิงทฤษฎีสามารถใช้อธิบายปรากฏการณ์ หลักการ หรือกฎเกณฑ์ต่างๆ ได้อย่างดีจนเป็นที่ยอมรับแล้ว จะเรียกแบบจำลองเชิงทฤษฎีนั้นว่าทฤษฎี

3. การนำความรู้และวิธีการทางวิทยาศาสตร์ไปใช้ (Application of Scientific Knowledge and methods) การสอนวิชาวิทยาศาสตร์ให้เกิดความรู้ความเข้าใจในเนื้อหาวิชาต่างๆ นั้น ควรได้ฝึกให้นักเรียนได้รู้จักนำความรู้และวิธีการต่างๆ ในวิชาวิทยาศาสตร์ไปใช้แก้ปัญหาใหม่ๆ ได้ด้วย

4. ทักษะปฏิบัติการใช้เครื่องมือ (Manual Skills) ในการเรียนการสอนวิทยาศาสตร์ ผู้เรียนต้องทำการทดลองเพื่อหาคำตอบของปัญหา จึงจำเป็นต้องฝึกให้ผู้เรียนได้มีทักษะในการใช้เครื่องมือวิทยาศาสตร์และทักษะในการติดตั้งเครื่องมือสำหรับทดลอง เพื่อให้เกิดความคล่องแคล่วในการปฏิบัติ ไม่ทำให้เครื่องมือที่ใช้ชำรุดเสียหายไม่เป็นอันตรายต่อตนเองและผู้อื่น

5. เจตคติและความสนใจ (Attitudes and Interests) เป็นการวัดพฤติกรรมด้านเจตพิสัย โดยเน้นการวัดความรู้สึก อารมณ์ การยอมรับ ซึ่งครูวิทยาศาสตร์ต้องการให้ผู้เรียนได้มีพัฒนาการเกี่ยวกับเจตคติและความสนใจวิทยาศาสตร์

6. การมีแนวโน้มในทางวิทยาศาสตร์ (Orientation) พฤติกรรมส่วนนี้เป็นการวัดหาความสัมพันธ์ระหว่างวิทยาศาสตร์กับเรื่องอื่นๆ การที่นักเรียนได้มองเห็นถึงความสัมพันธ์เหล่านี้จะช่วยให้ได้รับความก้าวหน้าทางวิทยาศาสตร์ได้เล็งเห็นคุณค่าและความสำคัญของวิทยาศาสตร์

ธงชัย ชิวปรีชา และคณะ (2526 : 298 – 326) ได้กล่าวถึงลักษณะของข้อสอบวัดพฤติกรรมด้านความรู้ความจำ พฤติกรรมด้านความเข้าใจ พฤติกรรมด้านกระบวนการสืบเสาะหาความรู้ และพฤติกรรมด้านการนำความรู้และวิธีการทางวิทยาศาสตร์ไปใช้ ซึ่งเป็นพฤติกรรมที่ครูต้องออกข้อสอบตามตารางวิเคราะห์เนื้อหาและพฤติกรรมไว้ดังต่อไปนี้

1. ลักษณะของข้อสอบวัดพฤติกรรมด้านความรู้ความจำ เป็นการวัดเกี่ยวกับเนื้อหา เรื่องราวหรือความรู้ต่างๆ ที่นักเรียนได้เรียนมาแล้ว

2. ลักษณะของข้อสอบวัดพฤติกรรมด้านความเข้าใจ สามารถวัดได้ 3 ลักษณะคือ

2.1 ลักษณะที่ 1 กำหนดสถานการณ์ใหม่มาให้แล้วให้นักเรียนระบุข้อเท็จจริง มโนคติ หลักการ หรือให้นักเรียนยกตัวอย่างสถานการณ์ใหม่ที่เกี่ยวข้องกับความรู้ที่กำหนดให้

2.2 ลักษณะที่ 2 กำหนดสถานการณ์ใหม่มาให้ แล้วให้นักเรียนยกตัวอย่างหรือระบุสถานการณ์อีกสถานการณ์หนึ่งที่เป็นไปตามวิธีการ หลักการ กฎ หรือทฤษฎีเดียวกัน

2.3 ลักษณะที่ 3 กำหนดสถานการณ์ใหม่มาให้ ซึ่งอาจอยู่ในรูปของข้อความ สัญลักษณ์รูปภาพ หรือแผนภาพ แล้วให้นักเรียนแปลความหมายของสถานการณ์ที่กำหนดให้

3. ลักษณะของข้อสอบวัดพฤติกรรมด้านกระบวนการสืบเสาะหาความรู้วิทยาศาสตร์ ประกอบด้วยพฤติกรรมย่อย 4 พฤติกรรมได้แก่

3.1 การสังเกตและการวัด สามารถออกข้อสอบเพื่อวัดพฤติกรรมด้านการสังเกตและการวัดในส่วนที่เกี่ยวกับการสังเกตวัตถุและปรากฏการณ์ต่างๆ การบรรยายผลการสังเกต การวัดขนาดของวัตถุรวมทั้งการเปลี่ยนแปลงวิธีการวัดในวิธีต่างๆ กัน การเลือกเครื่องมือที่เหมาะสม การรู้จักประเมินผลในการวัดและความถูกต้องของเครื่องมือได้

3.2 การมองเห็นปัญหาและการหาวิธีการแก้ปัญหา นั้น พฤติกรรมในกลุ่มนี้ได้แก่ การรับรู้ว่ามีปัญหานั้นเป็นอย่างไร การตั้งสมมติฐาน การเลือกวิธีการที่เหมาะสมในการทดสอบสมมติฐาน และการออกแบบการทดลองที่เหมาะสมสำหรับทดสอบสมมติฐาน

3.3 การแปลความหมายข้อมูลและการสร้างข้อสรุป ประกอบด้วยพฤติกรรมย่อย 6 พฤติกรรม ได้แก่ การจัดกระทำกับข้อมูล การนำเสนอข้อมูล การแปลความหมายของข้อมูล การทำนายทั้งแบบเพิ่มเติมและแบบขยายความ การตรวจสอบสมมติฐาน การสร้างข้อสรุป เป็นการเน้นการแปลความหมายข้อมูลที่ได้จากการทดลองทั้งที่อยู่ในรูปของการบรรยาย ตารางกราฟหรือแผนภาพต่างๆ

3.4 การสร้าง การทดสอบ และการปรับปรุงแบบจำลองเชิงทฤษฎี พฤติกรรมในกลุ่มนี้ที่อาจวัดได้ด้วยข้อสอบแบบเขียนตอบ ได้แก่ พฤติกรรมการระบุปรากฏการณ์และหลักการต่างๆ ที่อธิบายได้ด้วยแบบจำลองเชิงทฤษฎี การสร้างสมมติฐานจากแบบจำลองเชิงทฤษฎี การแปลความหมายและการประเมินผลการทดลองเพื่อตรวจสอบแบบจำลองเชิงทฤษฎี การปรับปรุงแก้ไขหรือเพิ่มเติมแบบจำลองทฤษฎี

4. ลักษณะของข้อสอบวัดพฤติกรรมด้านการนำความรู้และวิธีการวิทยาศาสตร์ไปใช้ จะมีลักษณะแบบยกสถานการณ์ใหม่หรือปัญหาใหม่ ซึ่งนักเรียนต้องใช้ความคิดหลายขั้นตอนมากกว่าข้อสอบวัดความเข้าใจ

โดยสรุปแล้ว การวัดผลสัมฤทธิ์ทางการเรียนในสถาบันการศึกษาต่าง ๆ นิยมสร้างแบบทดสอบวัดพฤติกรรมการเรียนรู้ตามทฤษฎีต่างๆ ที่มีผู้เสนอแนวคิดไว้ สำหรับวิชาวิทยาศาสตร์วัดอุปสงค์การเรียนรู้การสอนตามแนวคlopเฟอร์เป็นที่แพร่หลายมากในปัจจุบัน เนื่องจากมีความเหมาะสมกับการเรียนการสอนสามารถวัดได้ครอบคลุมทั้งเนื้อหาที่เป็นความรู้ทางวิทยาศาสตร์ และพฤติกรรมที่ต้องการให้เกิดขึ้นในตัวผู้เรียน ในการศึกษาครั้งนี้จึงสร้างแบบทดสอบโดยยึดแนวทฤษฎีของคlopเฟอร์

3. การจัดเรียงลำดับข้อคำถามของแบบทดสอบ

ทราเวอร์ (Travers. 1950 : 127 - 128) เสนอการจัดเรียงลำดับข้อคำถามไว้ 4 วิธี คือ

1. เรียงตามลำดับความยาก (Arrangement in Order of Difficulty) วิธีนี้เป็นวิธีที่ใช้เรียงลำดับข้อคำถามโดยทั่วไป ความยากของข้อคำถามจะพิจารณาโดยการคำนวณทางสถิติหรือจากการตัดสิน ข้อดีของการเรียงข้อสอบด้วยวิธีนี้ คือ ผู้สอบจะพบปัญหาง่ายในตอนแรก และไม่หมดกำลังใจเมื่อพบปัญหาที่ยากขึ้น แต่วิธีนี้มีข้อเสีย คือ เนื้อหาและพฤติกรรมที่ต้องการวัดไม่อยู่รวมเป็นพวกเดียวกัน ทำให้ผู้สอบต้องเปลี่ยนความคิดอย่างรวดเร็ว และข้อเสียอีกประการหนึ่ง คือ ผู้สอบอาจหมดกำลังใจที่จะทำแบบทดสอบ เพราะเมื่อเริ่มพบข้อยากแล้วข้อต่อไปของแบบทดสอบน่าจะยากยิ่งขึ้นไปอีก

2. เรียงลำดับความยากหมุนเวียน (Arrangement in Cyclic Order of Difficulty) วิธีนี้ มีจุดมุ่งหมายเพื่อกระตุ้นให้ผู้สอบอ่านข้อสอบทุกข้อ เพราะถ้าผู้สอบอ่านต่อไปอีกระยะก็จะพบข้อง่ายอีก การจัดแบบนี้มีประโยชน์แต่ผู้สอบต้องเปลี่ยนความคิดในแต่ละเรื่องอย่างรวดเร็ว การเรียงลำดับความยากแบบหมุนเวียนนี้ เพื่อหลีกเลี่ยงข้อเสีย ข้อที่สองของการเรียงลำดับข้อคำถามตามระดับความยาก

3. เรียงลำดับตามกลุ่มเนื้อหาวิชา (Arrangement According to Subject - Matter Area) ข้อดีของวิธีนี้คือผู้สอบจะคิดปัญหาในเนื้อหาเดียวกันก่อนที่จะเปลี่ยนไปคิดเนื้อหาอื่น ๆ

4. เรียงตามจุดมุ่งหมายของการวัด (Arrangement According to the Goals Measured) แบบทดสอบบางชุด ผู้จัดทำข้อสอบจะจัดเรียงเป็นกลุ่มตามพฤติกรรมการวัดประเภทเดียวกัน เช่น แบบทดสอบโค - ออฟเพอเรทีฟ (Co - operative Test) จะจัดข้อคำถามที่วัดความจำเกี่ยวกับคำศัพท์ และความคิดรวบยอด (Terms and Concepts) ไว้ด้วยกัน ส่วนการวัดความเข้าใจแยกไว้อีกพวกหนึ่ง

การเรียงข้อสอบแบบนี้มีข้อดีคือครูสามารถตรวจสอบว่าวัดตรงตามจุดมุ่งหมายที่ต้องการวัดหรือไม่ ส่วนข้อเสียที่พบได้ คือ ในทางปฏิบัติแล้ว จะไม่สามารถรวมกลุ่มของปัญหาจากเนื้อหาที่คล้ายกันเข้าไว้ด้วยกันได้ ✕

ลินควิสต์ (Lindquist. 1955 : 179) ให้ความเห็นเกี่ยวกับการจัดเรียงข้อคำถาม โดยเสนอแนะวิธีจัดเรียงข้อคำถามคล้ายกับ อนันต์ ศรีโสภะ และทราเวอร์ (Travers) ว่าในการเรียงลำดับข้อคำถามไม่จำเป็นต้องจัดเรียงลำดับข้อคำถามตามความยากง่าย ถ้าคำถามเหล่านั้นเป็นเนื้อหาเดียวกัน และมีความยากง่ายพอ ๆ กัน แต่ถ้าคำถามเหล่านั้นมีเนื้อหาต่างกัน และแต่ละข้อให้ 1 คะแนนต้องแยกข้อคำถามออกเป็นฉบับ (Subtests) และแยกเวลาในการสอบของแต่ละฉบับ การจัดเรียงข้อคำถามในแต่ละฉบับอาจขึ้นอยู่กับความสัมพันธ์ของเนื้อหาวิชา (Subject Matter) โดยให้สอบเนื้อหาทั่ว ๆ ไปก่อน แล้วค่อยสอบเนื้อหาที่เฉพาะในตอนหลัง

นันนาลลี (Nunnally. 1959 : 159) มีความเห็นว่า ไม่ว่าจะเป็นแบบทดสอบที่ใช้ในการทำนาย หรือวัดผลสัมฤทธิ์ทางการเรียน อาจเรียงลำดับข้อคำถามจากง่ายไปหายาก หรืออาจเรียงลำดับความยากง่ายสลับกันไปและข้อคำถามที่เรียงตามลำดับความยากนั้นจะต้องเป็นแบบทดสอบที่ทดลองสอบและหาความเชื่อมั่นมาแล้ว แต่ถ้ายังไม่ได้ทดสอบและปรับปรุงควรเรียงลำดับข้อคำถามแบบสุ่ม ซึ่งสแตนเลย์และฮอฟกินส์ (Stanley and Hopkins. 1972 : 191) ก็มีความเห็นในทำนองเดียวกันว่า การเรียงลำดับข้อคำถามควรเริ่มต้นจากข้อที่ง่ายที่สุดและให้ข้อยากที่สุดไว้ตอนท้ายของแบบทดสอบ เพราะถ้านำแบบทดสอบที่เป็นข้อยากๆ ไว้ตอนต้นๆ อาจทำให้เกิดความท้อใจได้ โดยเฉพาะสำหรับในกลุ่มนักเรียนที่มีความสามารถทางการเรียนระดับปานกลางและต่ำ

✕ ขวาล แพทย์กุล (2518 : 134) กล่าวว่าในข้อสอบแต่ละฉบับควรเรียงคำถามเริ่มจากข้อง่ายไปหาข้อยากเสมอ ไม่ควรเรียงคำถามตามลำดับการสอนก่อนหลัง ไม่ควรเรียงตามลำดับเนื้อหาวิชาที่สอน หรือเรียงตามความยากง่ายของบทเรียนเลยทีเดียว และถ้าเป็นข้อคำถามที่เริ่มใช้เป็นครั้งแรกต้องนึกคาดคะเนเอาเองว่าข้อใดยากกว่ากันแล้วเรียงลำดับตามนั้น

กรอนลันด์ (Gronlund. 1977 : 104 - 106) เสนอแนวคิดในการจัดเรียงข้อคำถามไว้ว่าควรจัดเรียงข้อคำถามประเภทเดียวกันไว้ด้วยกัน เช่น คำถามแบบอัตนัย คำถามแบบถูกผิด คำถามแบบเลือกตอบ และอื่น ๆ และในแต่ละประเภทควรจัดเรียงตามจุดมุ่งหมาย เช่น ความรู้ความจำ ความเข้าใจ การนำไปใช้ ฯลฯ โดยเรียงจากข้อที่ง่ายไปหาข้อที่ยาก

อนันต์ ศรีโสภะ (2520 : 141) กล่าวว่าแนวทางการจัดเรียงข้อสอบที่ดี ครูควรจัดเรียงจากข้อง่ายไปหาข้อยาก เพราะยังไม่มีหลักฐานยืนยันว่าการจัดเรียงข้อสอบที่มีเนื้อหาเดียวกันเข้าไว้ด้วยกันจะช่วยให้ค่าความเที่ยงตรง ค่าความเชื่อมั่นสูง

อุทุมพร จามรมาน (2536 : 54) เสนอหลักการจัดเรียงข้อคำถามในแบบทดสอบ ดังนี้

1. จัดข้อคำถามรูปแบบเดียวกันไว้ด้วยกันพร้อมทั้งเขียนคำชี้แจงในการตอบให้ผู้ตอบเข้าใจชัดเจน
2. จัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด
3. จัดเรียงข้อคำถามตามเนื้อหาของบทเรียน
4. จัดเรียงข้อคำถามตามจุดมุ่งหมายในการสอบ หรือพฤติกรรมที่จะต้องวัด
5. จัดเรียงข้อคำถามโดยการสุ่ม

ข้อดีของการจัดเรียงข้อคำถามในแต่ละรูปแบบของอุทุมพร จามรมาน

1. การจัดเรียงข้อคำถามตามรูปแบบข้อสอบจะทำให้ผู้สอบไม่สับสน ว่าควรตอบอย่างไรในข้อนั้น ๆ
2. การจัดเรียงข้อคำถามจากข้อง่ายไปหาข้อยาก จะทำให้ผู้สอบมีกำลังใจในการทำแบบทดสอบ
3. การจัดเรียงข้อคำถามตามเนื้อหาและจุดมุ่งหมาย หรือพฤติกรรมที่ต้องการวัดจะทำให้ง่ายต่อการวิเคราะห์ว่าผู้สอบมีความรู้ในเนื้อหาอะไรบ้าง บรรลุตามจุดมุ่งหมายใดบ้าง หรือมีพฤติกรรมการเรียนรู้ตามที่กำหนดไว้ในด้านใดบ้าง

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้สนใจศึกษาการจัดเรียงข้อคำถามตามค่าความยากที่เรียงจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

4. ความเที่ยงตรงของแบบทดสอบ

ความเที่ยงตรงของแบบทดสอบเป็นคุณสมบัติหนึ่งที่สำคัญ มีนักวัดผลการศึกษาได้ให้ความหมายไว้ดังนี้

อนาสตาซี (Anastasi. 1968 : 29) ให้ความหมายของความเที่ยงตรงของแบบทดสอบไว้ว่า เป็นระดับที่แทนแบบทดสอบวัดได้จริงในสิ่งที่ต้องการจะวัด

สแตนเลย์ (Stanley. 1972 : 101) ได้กล่าวว่า ความเที่ยงตรงของการวัด หมายถึง การวัดนั้นตรงตามหน้าที่ที่จะวัดได้ดีเพียงใด หรือระดับที่การวัดสามารถบรรลุจุดมุ่งหมายบางอย่าง

ทักแมน (Tuckman. 1975 : 229) กล่าวว่า ความเที่ยงตรงของแบบทดสอบ หมายถึง แบบทดสอบฉบับหนึ่งวัดในสิ่งที่เราต้องการให้วัดได้หรือไม่

วรรณดี ม้าลำพอง (2530 : 147) ได้กล่าวว่า ความเที่ยงตรง หมายถึง ความสามารถในการบอกค่าที่ต้องการวัดได้ถูกต้อง

มิ่ง เทพครเมือง (2539 : 15) กล่าวว่า ความเที่ยงตรงของแบบทดสอบ หมายถึง คุณสมบัติของแบบทดสอบ หรือเครื่องมือวัดใด ๆ ที่สามารถวัดสิ่งที่ต้องการวัดได้อย่างถูกต้อง ตรงตามจุดมุ่งหมายที่ต้องการจะวัด

อัมพร วิชัยศรี (2541 : 17) ให้ความหมายไว้ว่า ความเที่ยงตรงของแบบทดสอบ หมายถึง แบบทดสอบหรือเครื่องมือวัดใด ๆ ที่วัดได้ตรงตามจุดมุ่งหมายที่ต้องการวัด

จากความหมายของความเที่ยงตรงที่บุคคลต่าง ๆ กล่าวมา พอจะสรุปได้ว่า ความเที่ยงตรงของแบบทดสอบ หมายถึง คุณสมบัติของเครื่องมือวัด หรือแบบทดสอบใด ๆ ที่สามารถวัดได้ตรงตามจุดมุ่งหมายหรือหน้าที่ที่จะวัดได้ดีและเป็นระดับที่แทนของเครื่องมือวัดในสิ่งที่ต้องการวัดได้จริง ดังนั้นการพิจารณาความเที่ยงตรง จึงพิจารณาที่ความเที่ยงตรงของการแปลความหมายของคะแนนโดยเป็นการมุ่งตอบคำถาม 2 ประการ คือ

1. คะแนนจากการสอบ สามารถพยากรณ์ถึงความสามารถตามเกณฑ์ที่ต้องการได้ถูกต้องแม่นยำหรือไม่เพียงพอ ตรวจสอบโดยหาความสัมพันธ์ระหว่างคะแนนจากการสอบกับคะแนนตามเกณฑ์

2. แบบตรวจสอบมุ่งวัดเนื้อหาและพฤติกรรมใด เป็นการตรวจสอบถึงลักษณะความสามารถ และลักษณะพฤติกรรมที่เป็นกลุ่มตัวอย่างในแบบทดสอบ

ชนิดของความเที่ยงตรง

อนันต์ ศรีโสภ (2524 : 69) แบ่งความเที่ยงตรงตามหลักฐานที่นำมาแสดงออกเป็น 3 แบบคือ

1. ความเที่ยงตรงตามเนื้อหา (Content Validity)
2. ความเที่ยงตรงเชิงสัมพันธ์กับเกณฑ์ (Criterion – related Validity)
3. ความเที่ยงตรงเชิงโครงสร้าง (Construct Validity)

ความเที่ยงตรงตามเนื้อหาและการตรวจสอบ

สำริง บุญเรืองรัตน์ (2529 : 90) กล่าวว่า iva ความเที่ยงตรงตามเนื้อหา หมายถึง คุณภาพของแบบทดสอบที่สร้างคำถามให้ได้สัดส่วน หรือนำหนักความสำคัญของเนื้อหาแต่ละเรื่องที่จะวัด

การตรวจสอบความเที่ยงตรงตามเนื้อหาของแบบทดสอบจะกระทำด้วยการวิเคราะห์เชิงเหตุผล โดยอาศัยผู้เชี่ยวชาญด้านเนื้อหา (Subject – Matter Specialist) โดยตรง ซึ่งการหาความเที่ยงตรงตามเนื้อหาเป็นการพิจารณาในลักษณะรวม ๆ ไม่ได้ชี้ในลักษณะเชิงปริมาณ หรือตัวเลข ซึ่งวิธีการประเมินความเที่ยงตรงตามเนื้อหา อูทุมพร จามรมา (2532 : 24-28) กล่าวว่า iva 2 ระยะ ดังนี้

ระยะแรก เป็นการตรวจสอบขั้นตอนการสร้างโดยผู้เชี่ยวชาญอย่างน้อย 3 ท่าน พิจารณาเนื้อหาและพฤติกรรมที่ต้องการวัดที่กำหนดขึ้นเกี่ยวกับ 1. เนื้อหาที่จำแนกแตกย่อยออกเป็นหมวดหมู่ หรือหัวข้อใหญ่ หัวข้อย่อย ถูกต้อง ครบถ้วนสมบูรณ์ตามหลัก กฎ ทฤษฎีของเนื้อหาที่ต้องการวัดหรือไม่ 2. น้ำหนักของหัวข้อใหญ่ หัวข้อย่อยของแต่ละเนื้อหา และแต่ละพฤติกรรมที่ต้องการวัดมีความถูกต้อง เหมาะสมหรือไม่ 3. ข้อสอบแต่ละข้อ ถามได้ตรงหรือสอดคล้องกับเนื้อหา และพฤติกรรมที่ต้องการวัดหรือไม่

ระยะหลัง เป็นการตรวจสอบหลังจากร่างแบบทดสอบเสร็จเรียบร้อยแล้ว แล้วนำไปให้ผู้เชี่ยวชาญหลายๆ ท่านที่มีความเห็นตรงกันในประเด็นเกี่ยวกับ 1. ความครอบคลุมของเนื้อหา 2. ความสอดคล้องระหว่างเนื้อหาที่แยกเป็นหัวข้อใหญ่กับหัวข้อย่อย ๆ 3. ความสอดคล้องระหว่างเนื้อหา พฤติกรรมและน้ำหนัก

ความเที่ยงตรงเชิงสัมพันธ์กับเกณฑ์และการตรวจสอบ

ความเที่ยงตรงเชิงสัมพันธ์กับเกณฑ์ หมายถึง ความเที่ยงตรงที่ได้จากความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบนั้นกับเกณฑ์ภายนอก (External Criterion) (สำริง บุญเรืองรัตน์. 2529 : 90-91 ; ต่าย เชียงฉี. 2526 : 146-148 ; อรุณ ศรีสะอาด. 2526 : 193 ; บุญธรรม กิจปรีดาบริสุทธิ์. 2535 : 155) ได้แบ่งออกเป็น 2 ประเภท คือ

1. ความเที่ยงตรงตามสภาพ (Concurrent Validity) หมายถึง คุณภาพของแบบทดสอบที่สามารถบ่งชี้ได้อย่างถูกต้อง ตามสภาพที่เป็นจริงของผู้ถูกสอบ เช่น ถ้าแบบทดสอบใดบ่งชี้ว่าผู้ถูกสอบเป็นผู้ที่มีความสามารถวเลขได้คล่องแคล่ว และถูกต้องเป็นจำนวนมาก ในชีวิตจริงเมื่อเด็กคนนี้ไปซื้อของก็สามารถคิดราคาของ หรือคิดเงินทอนได้รวดเร็ว เป็นต้น

การหาความเที่ยงตรงตามสภาพนี้ทำได้โดยหาค่าสัมประสิทธิ์ระหว่างคะแนนจากแบบทดสอบนั้นกับคะแนนการจัดอันดับของสภาพที่เป็นจริงของนักเรียนในกลุ่มนั้น ซึ่งอาจใช้สูตรของ Person Product

Moment หรือ Spearrman Rank-Order ถ้าค่าสัมประสิทธิ์สหสัมพันธ์สูง แสดงว่าแบบทดสอบนั้นมีความเที่ยงตรงตามสภาพสูง

2. ความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) หมายถึง คุณภาพในการวัดที่สามารถนำผลการวัดนั้นไปพยากรณ์ความสำเร็จในการเรียนวิชาต่าง ๆ อาชีพต่าง ๆ ในอนาคตได้อย่างถูกต้อง

การหาความเที่ยงตรงเชิงพยากรณ์นี้หาได้โดยการหาสหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบนั้นกับผลการเรียน หรือความสำเร็จในการทำงานในอนาคต โดยใช้สูตรของ Person Product Moment ถ้าค่าสหสัมพันธ์สูง แสดงว่ามีความเที่ยงตรงเชิงพยากรณ์สูง

ความเที่ยงตรงตามโครงสร้างและการตรวจสอบ

สำเร็จ บุญเรืองรัตน์ (2529 : 90) กล่าวว่าไว้ว่า ความเที่ยงตรงเชิงโครงสร้าง หมายถึง คุณภาพของแบบทดสอบที่วัดได้ตรงตามโครงสร้างของพฤติกรรม หรือสมรรถภาพที่ต้องการวัด และวรรณคดี มัลลาลพอง (2530 : 152 - 154) กล่าวว่าไว้ว่าการหาความเที่ยงตรงเชิงโครงสร้าง มีวิธีการหาได้หลายวิธี ดังนี้

1. ใช้ความแตกต่างของความสามารถของต่างอายุเป็นเกณฑ์ (Age Differentiation) คือ การตรวจสอบความแตกต่างของคะแนนระหว่างผู้สอบที่มีอายุต่างกัน ซึ่งตรวจสอบการสร้างแบบทดสอบตามทฤษฎีพื้นฐานที่ว่า สถิติปัญญาจะเพิ่มขึ้นตามอายุ ทำได้โดยนำแบบทดสอบไปสอบกับเด็กที่มีอายุต่างกัน ถ้าผลการสอบพบว่าเด็กที่มีอายุสูงขึ้น คะแนนสอบเพิ่มขึ้น แสดงว่า แบบทดสอบฉบับนั้นมีความเที่ยงตรงเชิงโครงสร้าง

2. โดยหาความสัมพันธ์กับแบบทดสอบฉบับอื่น (Correlations with other tests) ทำได้โดยนำแบบทดสอบที่หาความเที่ยงตรงตามโครงสร้างกับแบบทดสอบที่วัดในคุณลักษณะเดียวกันที่เป็นมาตรฐานไปสอบกับกลุ่มตัวอย่างเดียวกัน ก็จะได้ข้อมูล 2 ชุด แล้วหาค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลทั้งสอง ก็จะได้ความเที่ยงตรงตามโครงสร้างตามต้องการ

3. วิธีหาความคงที่ภายใน (Internal Consistency) วิธีนี้ใช้คะแนนรวมของการทำแบบทดสอบเป็นเกณฑ์ แบบทดสอบที่ประกอบด้วยแบบทดสอบย่อย ๆ หลายตอนสามารถตรวจสอบความเที่ยงตรงตามโครงสร้าง โดยหาสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบทดสอบย่อยกับคะแนนรวมทั้งฉบับ ถ้าแบบทดสอบย่อยฉบับใดมีค่าสหสัมพันธ์ กับแบบทดสอบทั้งสองฉบับต่ำ จะพิจารณาตัดออกไปเพื่อให้แบบทดสอบฉบับนั้นมีความเที่ยงตรงตามโครงสร้าง

4. เปรียบเทียบคะแนนการสอบก่อน และ หลังการทดลอง (Pretest - Posttest Technique) โดยทดลองสอน หรือฝึกคุณลักษณะ หรือความสามารถตามจุดมุ่งหมายของการวัดของแบบทดสอบกับกลุ่มตัวอย่างหนึ่ง แล้วตรวจสอบดูว่าหลังการฝึก กลุ่มตัวอย่างมีความก้าวหน้าเพิ่มขึ้นหรือไม่ โดยเปรียบเทียบกับคะแนนหลังฝึกอบรมกับก่อนฝึกอบรม

5. วิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) วิธีการนี้เป็นวิธีการทางสถิติสำหรับตรวจสอบคุณสมบัติทางจิตวิทยา ด้วยการวิเคราะห์หาความสัมพันธ์ระหว่างคำถามแต่ละข้อเพื่อระบุลักษณะร่วมของคำถามเหล่านั้น ว่าข้อคำถามที่วัดนั้นประกอบด้วยองค์ประกอบอะไรบ้าง ตรงตามทฤษฎีหรือข้อสมมติฐานที่ตั้งไว้หรือไม่ ถ้าตรงก็แสดงว่ามีความเที่ยงตรงเชิงโครงสร้าง

6. วิธี (Multitrait Multimethod) เป็นวิธีที่เสนอโดยแคมป์เบลล์ และฟิสต์ (Lemke. 1976 ; citing Campbell and Fiske ปี ค.ศ. 1959) แคมป์เบลล์ชี้ให้เห็นว่าการหาความเที่ยงตรงตามโครงสร้างนั้นไม่ควรแสดงเพียงแต่ว่า แบบทดสอบนั้นมีค่าสหสัมพันธ์กันตามทฤษฎีเท่านั้น (Convergent Validity Coefficient)

แต่ต้องแสดงให้เห็นด้วยว่า ไม่ได้ให้ค่าสัมพัทธ์กันอย่างน้อยสำคัญกับตัวแปรอื่นที่แตกต่างกัน (Divergent Validity Coefficient)

ความเที่ยงตรงแบบลักษณะหลากหลาย วิธีหลาย (Multitrait Multimethod)

นักวัดผลหลายท่านได้อธิบายความหมายของการวิเคราะห์ความเที่ยงตรงแบบลักษณะหลากหลายวิธีหลาย (Multitrait Multimethod) ว่าเป็นความเที่ยงตรงเชิงโครงสร้างชนิดหนึ่งซึ่งพัฒนาขึ้นโดย Campbell และ Fiske ปี ค.ศ. 1959 (เบญจวรรณ แยมละมูล. 2538 : 21 - 23 ; Allen and Yen. 1979 : 109 - 115.) วิธีการหาความเที่ยงตรงแบบนี้ใช้กับแบบทดสอบที่มีหลายลักษณะ (Traits) และใช้วัดด้วยหลายวิธีการ (Method) ตัวอย่างเช่น แบบวัดลักษณะการเก็บตัวและอาการทางประสาท โดยใช้วิธีวัดด้วยวิธีการแบบถูกผิด และหลายตัวเลือก ดังนั้นจะได้แบบวัดทั้งหมด 4 ฉบับ นำแบบวัดนี้ไปทดสอบกับกลุ่มตัวอย่าง เมื่อนำมาวิเคราะห์จะได้ผล ดังตารางเมทริกซ์ความเที่ยงตรงแบบ (Multitrait Multimethod) ดังตาราง 6

ตาราง 6 ตารางเมทริกซ์ของความเที่ยงตรงแบบ (Multitrait Multimethod)

แบบวัดลักษณะ (ชนิดการวัด)	การเก็บตัว (ถูก - ผิด)	อาการทางประสาท (ถูก - ผิด)	การเก็บตัว (หลายตัวเลือก)	อาการทางประสาท (หลายตัวเลือก)
การเก็บตัว (ถูก - ผิด)	.80	.20	.75	.17
อาการทางประสาท (ถูก - ผิด)	.20	.85	.15	.71
การเก็บตัว (หลายตัวเลือก)	.75	.15	.83	.21
อาการทางประสาท (หลายตัวเลือก)	.17	.71	.21	.91

จากค่าสหสัมพันธ์ในตารางจะพบว่า มีค่าที่อยู่ในความหมายของ ค่าความเชื่อมั่น (Reliability) เป็นค่า .80, .85, .83 และ .91 และค่าที่วัดลักษณะเดียวกันแต่ต่างวิธีกัน มีค่าสหสัมพันธ์เท่ากับ .75 และ .71 ค่าเหล่านี้คือค่าที่อยู่ในความหมายของความเที่ยงตรงเชิงเหมือน ซึ่งมีค่าสหสัมพันธ์กันสูง ส่วนค่าที่อยู่ในความหมายความเที่ยงตรงเชิงจำแนกนั้น คือค่าสหสัมพันธ์ที่ได้จากการวัดลักษณะต่างกัน ถึงแม้จะใช้วิธีวัดแบบเดียวกัน หรือต่างวิธีกันจะมีค่าสหสัมพันธ์กันต่ำ ซึ่งได้แก่ค่า .15, .17, .20 และ .21 ดังแสดงความหมายในตาราง 7 ต่อไปนี้

ตาราง 7 ตารางเมทริกซ์ของความเที่ยงตรงแบบ (Multitrait Multimethod) ที่วัด 2 ลักษณะ (Traits) และใช้ วิธีวัด 2 วิธี (Methods)

แบบวัด	A ₁	B ₂	A ₂	B ₂
A ₁	สูง	ต่ำ	สูง	ต่ำ
B ₂	ต่ำ	สูง	ต่ำ	สูง
A ₂	สูง	ต่ำ	สูง	ต่ำ
B ₂	ต่ำ	สูง	ต่ำ	สูง

ในตารางความหมายคำว่า “สูง” แสดงถึง ค่าความเที่ยงตรงเชิงเหมือน ส่วนความหมายคำว่า “ต่ำ” แสดงถึงความเที่ยงตรงเชิงจำแนก

นอกจากนี้ความเที่ยงตรงแบบลักษณะหลากหลาย วิธีหลาย นั้นแคมป์เบลและฟิสต์ (Campbell and Fiske, 1959 ; Crocker and Algina, 1986 : 232 - 233) ได้อธิบายว่าเป็นการหาค่าสหสัมพันธ์ระหว่างลักษณะที่วัด และวิธีวัดของแบบทดสอบโดยแบ่งเป็น 3 แบบดังนี้

1. ค่าสหสัมพันธ์ที่อยู่ในรูปความเชื่อมั่นของแบบทดสอบ
2. ค่าสหสัมพันธ์ที่เป็นค่าสัมประสิทธิ์ความเที่ยงตรงเชิงเหมือน ซึ่งเป็นค่าสหสัมพันธ์ที่เกิดจากการวัดลักษณะเดียวกัน วิธีเดียวกัน หรือต่างวิธีกันก็ตามจะมีค่าสหสัมพันธ์สูง
3. ค่าสหสัมพันธ์ที่เป็นค่าสัมประสิทธิ์ความเที่ยงตรงเชิงจำแนก ซึ่งเป็นค่าสหสัมพันธ์ที่วัดลักษณะต่างกันถึงแม้จะใช้วิธีวัดแบบเดียวกันหรือต่างวิธีกันจะมีค่าสหสัมพันธ์ต่ำกว่าค่าสหสัมพันธ์ของความเที่ยงตรงเชิงเหมือน

โดยสรุปแล้ว ความเที่ยงตรงแบบลักษณะหลากหลาย วิธีหลาย (Multitrait Multimethod Validity) เป็นความเที่ยงตรงประเภทหนึ่งของความเที่ยงตรงเชิงโครงสร้าง (Construction Validity) ซึ่งมีความคิดพื้นฐานมาจากการตรวจสอบกระบวนการของค่าสหสัมพันธ์ ระหว่างแบบวัดทั้งหลายที่วัดลักษณะต่าง ๆ (Traits) ด้วยวิธีที่ต่างกัน แบบวัดต่าง ๆ ที่วัดลักษณะเดียวกันควรมีสหสัมพันธ์สูงมาก ในทางตรงกันข้ามแบบวัดต่าง ๆ ที่วัดลักษณะที่ต่างกันควรมีสหสัมพันธ์ที่ต่ำ และนอกจากนี้ยังไม่ควรมีอคติอันเนื่องมาจากการใช้วิธีที่แตกต่างกัน

5. ความเชื่อมั่นของแบบทดสอบ (Reliability)

5.1 ความหมายของความเชื่อมั่น

มีนักการศึกษา และนักจิตวิทยาหลายท่านได้ให้ความหมายของความเชื่อมั่นไว้ต่าง ๆ กัน ดังเช่น นันนาลลี (Nunnally, 1964 : 59) ให้ความหมายความเชื่อมั่นว่า เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการสอบ

ลินด์วอลล์และนิตโค (Lindvall and Nitko, 1967 : 126) ให้ความหมายความเชื่อมั่นว่าเป็นค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบสองครั้ง โดยใช้แบบทดสอบฉบับเดียวกันในเวลาต่างกัน

อนาสตาซี (Anastasi. 1968 : 71) ให้ความหมายความเชื่อมั่นว่าเป็นค่าคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันในเวลาที่แตกต่างกัน

ลอร์ด และโนวิก (Lord and Novick. 1968 : 198) ให้ความหมายความเชื่อมั่นว่าเป็นความคงที่ของคะแนนจากการตอบแบบทดสอบซ้ำ และคะแนนที่ได้จากการตอบแบบทดสอบทั้งสองครั้งที่เป็อิสระไม่ขึ้นกับความคลาดเคลื่อนของการวัดใด ๆ

กรอนลันด์ (Gronlund. 1976 : 93) ให้ความหมายความเชื่อมั่นว่าเป็นความคงที่จากการประเมินจากการวัดครั้งแรก และครั้งอื่น ๆ หรือค่าคงที่ของคะแนนในการทดสอบ

ชวาล แพร์ตกุล (2516 : 136) สำเรัง บุญเรืองรัตน์ (2529 : 38) ให้ความหมายความเชื่อมั่นว่าเป็นความคงที่ของคะแนนที่ได้จากการทดสอบ แบบทดสอบที่มีความเชื่อมั่นสูงไม่ว่าจะนำไปวัดกับบุคคลเดียวกันกี่ครั้งก็ตามก็จะได้ค่าคงที่เช่นเดิม ในทางตรงกันข้ามแบบทดสอบที่มีค่าความเชื่อมั่นต่ำ ค่าที่ได้จากการทดสอบสองครั้งของบุคคลกลุ่มเดียวกันจะมีการเปลี่ยนแปลงมาก

บุญเชิด ภิญญอนันตพงษ์ (2521 : 269) ให้ความหมายความเชื่อมั่นว่าเป็นความคงที่แน่นอนของคะแนนที่ได้จากการสอบวัดนักเรียนกลุ่มเดียวกัน ด้วยแบบทดสอบฉบับเดียวกันหลาย ๆ ครั้ง หรือด้วยแบบทดสอบสองฉบับที่มีลักษณะเสมอเหมือนกันภายใต้เงื่อนไขของตัวแปรอื่น ๆ ในการวัดนั้น ๆ

ล้วน สายยศและอังคณา สายยศ (2539 : 209) ให้ความหมายความเชื่อมั่นว่าหมายถึงความคงที่ของคะแนนที่ได้จากการสอบนักเรียนคนเดียวกันหลาย ๆ ครั้ง ในแบบทดสอบชุดเดิม

จากแนวคิดของนักการศึกษาและนักจิตวิทยาพอสรุปได้ว่าความเชื่อมั่นเป็นความคงที่ของคะแนนที่ได้จากการทดสอบ แบบทดสอบที่มีความเชื่อมั่นสูง ไม่ว่าจะนำไปวัดกับบุคคลเดียวกันกี่ครั้งก็ตาม ก็จะได้ค่าคงที่เช่นเดิม ส่วนแบบทดสอบที่มีค่าความเชื่อมั่นต่ำ ค่าที่ได้จากการสอบสองครั้งของบุคคลกลุ่มเดียวกันจะมีการเปลี่ยนแปลงมาก

5.2 การประมาณค่าความเชื่อมั่น

เมห์เรนส์ และเลห์แมน (พิมพร ภูมิรา. 2544 : 37 อ้างอิงจาก Mehrens and Lehmann, 1984 : 271 - 272) กล่าวถึงวิธีการประมาณค่าความเชื่อมั่น ดังนี้

1. วิธีวัดความคงตัว (Measures of Stability)
2. วิธีวัดความสมมูล (Measures of Equivalence)
3. วิธีวัดความสมมูล และความคงตัว (Measures of Equivalence and Stability)
4. วิธีวัดความสอดคล้องภายใน (Measures of Internal - Consistency)

4.1 วิธีแบ่งครึ่งแบบทดสอบ (Split - half)

4.2 วิธีของคูเดอร์ - ริชาร์ดสัน (Kuder - Richardson Estimates)

4.3 วิธีสัมประสิทธิ์อัลฟ่า (Coefficient Alpha)

4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt 's Analysis of Variance Procedure)

5. ความเชื่อมั่นของผู้ให้คะแนน (Score Judge Reliability)

อนันต์ ศรีโสภา (2524 : 49 - 57) กล่าวถึงการคำนวณค่าความเชื่อมั่นว่ามีวิธีต่าง ๆ หลายวิธี ซึ่งในแต่ละวิธีทำให้เกิดความคลาดเคลื่อนไม่เหมือนกันการที่จะเลือกใช้วิธีใดนั้นขึ้นอยู่กับความมุ่งหมายและลักษณะของคะแนนที่ได้จากการทดสอบนั้น วิธีหาค่าความเชื่อมั่นโดยทั่ว ๆ ไปมี 4 แบบ คือ

1. แบบสอบซ้ำ (Measures of Stability)

เป็นการวัดความคงที่ของแบบทดสอบ โดยนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มเดียวกันสองครั้งในเวลาที่ยาวนาน แล้วนำผลการทดสอบทั้งสองมาหาค่าสัมประสิทธิ์สหสัมพันธ์โดยวิธี Pearson Product Moment Correlation การหาค่าความเชื่อมั่นโดยวิธีนี้มีความคลาดเคลื่อนที่เห็นได้ชัด คือ ช่วงระยะเวลาที่ทดสอบครั้งแรกและทดสอบครั้งหลัง ความแตกต่างของช่วงระยะเวลาทดสอบนี้นักเรียนจะอาศัยความจำจากการสอบครั้งแรก ทำให้คะแนนสอบครั้งหลังสูงขึ้นและ มีค่าความเชื่อมั่นสูง

2. แบบใช้ข้อสอบคล้ายกัน (Measures of Equivalence)

เป็นการหาความเชื่อมั่นโดยใช้แบบทดสอบสองชุดที่มีเนื้อหา ค่าเฉลี่ย และค่าความแปรปรวนเท่ากันโดยทดสอบกับนักเรียนกลุ่มเดียวกันในวันเดียวกันแล้วนำผลการสอบทั้งสองฉบับนั้นมาหาค่าสัมประสิทธิ์สหสัมพันธ์ หรือค่าความเชื่อมั่น การหาค่าความเชื่อมั่นวิธีนี้มีความคลาดเคลื่อนเนื่องจากความแตกต่างของสเกลหรือความแตกต่างของข้อสอบสองฉบับ การที่จะสร้างให้มีค่าเฉลี่ยและความแปรปรวนเท่ากันนั้นทำได้ยากมากในทางปฏิบัติ

3. แบบใช้ข้อสอบคล้ายกัน และสอบซ้ำ (Measures of Equivalence and Stability)

เป็นการหาความเชื่อมั่นโดยนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มเดียวกันในเวลาที่ยาวนาน แล้วนำคะแนนผลการสอบทั้งสองมาหาค่าสัมประสิทธิ์สหสัมพันธ์ สาเหตุของการเปลี่ยนแปลงคะแนนจากวิธีนี้อาจเนื่องมาจากความไม่คงที่ของนักเรียนและตัวข้อสอบค่าความเชื่อมั่นที่ได้โดยวิธีนี้จะมีค่าต่ำกว่าสองวิธีแรก

4. แบบวัดความคงที่ภายใน (Measures of Internal Consistency)

เป็นการหาค่าความเชื่อมั่น โดยนำแบบทดสอบไปทดสอบกับนักเรียนเพียงครั้งเดียวแล้วนำไปคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแต่ละข้อกับคะแนนรวม การหาค่าความเชื่อมั่นวิธีนี้แบ่งออกเป็น 4 แบบ คือ

4.1 แบบแบ่งครึ่งจำนวนข้อสอบ (Split - half)

เป็นวิธีคำนวณค่าความเชื่อมั่นของแบบทดสอบโดยนำแบบทดสอบฉบับเดียวไปทดสอบกับนักเรียนกลุ่มหนึ่งครั้งเดียว แล้วแบ่งแบบทดสอบออกเป็นสองชุด นำคะแนนของทั้งสองชุดมาหาค่าสัมประสิทธิ์ r_{tt} คือค่าความเชื่อมั่นของแบบทดสอบฉบับย่อยที่แยกไป หรือเป็นค่าความเชื่อมั่นของแบบทดสอบครึ่งหนึ่งของแบบทดสอบฉบับเดิม ถ้าต้องการทราบค่าความเชื่อมั่นทั้งฉบับก็ปรับขยายด้วยสูตรของ Spearman Brown ค่าที่ได้คือค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4.2 Kuder - Richardson Estimates

เป็นการหาค่าความเชื่อมั่นโดยใช้สูตรของ Kuder - Richardson ประกอบด้วย KR - 20 และ KR - 21 การคำนวณค่าความเชื่อมั่นโดยใช้สูตร KR - 21 จะได้ค่าต่ำกว่าการคำนวณโดยใช้สูตร KR - 20 เนื่องจาก KR - 21 ใช้ค่าเฉลี่ย (Mean) ของข้อสอบทุกข้อแทนความแปรปรวนของข้อสอบแต่ละข้อการหาความเชื่อมั่นวิธีนี้เพื่อแก้ปัญหาการ Split แบบทดสอบสองฉบับย่อยให้ Equivalence Form กันได้

4.3 Hoyt's Analysis of Variance Procedure

เป็นการหาค่าความเชื่อมั่นของแบบทดสอบโดยการวิเคราะห์ค่าความแปรปรวน (Analysis of Variance) หรือเรียกว่า ANOVA วิธีนี้ให้ค่าความเชื่อมั่นเหมือนกับ KR - 20

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

1. การวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ของการวัดหลาย ๆ ครั้ง จากการสอบซ้ำด้วยแบบทดสอบฟอร์มเดียวกัน หรือ ฟอร์มที่คู่ขนานกัน ซึ่งมีค่าเท่ากับความแปรปรวนของคะแนนจริงหารด้วยความแปรปรวนของคะแนนที่สอบได้ โดยเรียกสัมประสิทธิ์สหสัมพันธ์ที่คำนวณจากคะแนนที่ได้จากการสอบด้วยข้อสอบฉบับเดียวกันสองครั้งในเวลาที่ต่างกันว่า สัมประสิทธิ์ของความคงตัว (Coefficient of Stability) และเรียกสัมประสิทธิ์สหสัมพันธ์ที่คำนวณจากคะแนนที่ได้จากการสอบด้วยข้อสอบคู่ขนาน (Parallel test) หรือแบบสลับฟอร์ม (Alternate Form) ที่สอบติดต่อกันทันที หรือสอบทั้งช่วงเวลาว่าสัมประสิทธิ์ของความสมมูล (Coefficient of Equivalent)

2. การวิเคราะห์สัมประสิทธิ์ความสอดคล้องภายในของแบบทดสอบ วิธีนี้จะใช้แบบทดสอบฉบับเดียวและสอบครั้งเดียว แล้วคำนวณหาสัมประสิทธิ์ความสอดคล้องภายในของแบบทดสอบ (Coefficient of Internal Consistency) จากความแปรปรวนของคะแนนสอบภายในฉบับจากส่วนย่อยแต่ละตอนและส่วนรวมทั้งฉบับ

นักทฤษฎีทางการวัดผล ได้เสนอเทคนิคในการประมาณค่าความเชื่อมั่นที่เชื่อถือได้ไว้หลายเทคนิค โดยมีข้อสันนิษฐานเบื้องต้น (Presumptions) ว่าแบบทดสอบฉบับรวมสามารถแบ่งเป็นส่วน ๆ เช่น สองส่วน สามส่วน สี่ส่วน หรือหลาย ๆ ส่วน และเมื่อใช้ระดับของความคู่ขนานของการวัด (Degree of Measurement Parallelism) ในแต่ละส่วนเป็นเกณฑ์แล้ว การประมาณค่าความเชื่อมั่นที่ได้จากการสอบเพียงครั้งเดียวด้วยข้อสอบเพียงฉบับเดียว สามารถจัดกลุ่มตามข้อตกลงของระดับความคู่ขนานได้ 3 กลุ่ม (บุญเชิด ภิญโญนนตพงษ์. 2537 : 9 - 46) ดังนี้

1. แบบจำลองความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts)

เป็นการประมาณค่าความเชื่อมั่นจากแบบทดสอบแต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิมที่มีข้อตกลงอย่างเคร่งครัด 6 ข้อ คือ

1.1 มีความเป็นเอกพันธ์ในเนื้อหา หรือวัดคุณลักษณะเดียวกัน

1.2 มีคะแนนจริงเท่ากัน ($T_{x1} = T_{x2} = T_{x3} = \dots$) มีความแปรปรวนคลาดเคลื่อนเท่ากัน

$$(S^2_{E1} = S^2_{E2} = S^2_{E3} = \dots)$$

1.3 มีคะแนนสอบ (X) เฉลี่ยเท่ากัน ($\mu_1 = \mu_2 = \mu_3 = \dots$)

1.4 มีความแปรปรวนของคะแนนสอบ (X) เท่ากัน ($S^2_{x1} = S^2_{x2} = S^2_{x3} = \dots$)

1.5 มีความแปรปรวนร่วมของคะแนนสอบ (X) กับคะแนนสอบส่วนอื่น ๆ เท่ากัน

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

1.6 มีความแปรปรวนร่วมของคะแนนสอบ (X) กับคะแนนเกณฑ์ภายนอกเท่ากัน

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

2. แบบจำลองคะแนนจริงจำเป็นต้องสมมูล (Essentially Tau - Equivalent Parts)

วิธีนี้เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบ โดยผ่อนปรนเงื่อนไขข้อ 2, 3 และ 4 ของแบบจำลองความคู่ขนานแบบมาตรฐานเดิมให้มีความเป็นไปได้มากขึ้น เนื่องจากในทางปฏิบัติแทบไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานตามแบบมาตรฐานเดิมได้ จึงได้มีการพัฒนาเทคนิคที่เหมาะสมขึ้นมาใหม่ ดังนี้

2.1 คะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากันแต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วน นั่นคือคะแนนจริงส่วนที่หนึ่งเท่ากับคะแนนจริงส่วนที่สองรวมกับค่าคงที่ค่าหนึ่ง ดังนี้

$$T_{ig} = T_{ih} + C_{gh} \text{ เมื่อ } g = h = 1, \dots, k \text{ และ } C_{gh} \text{ ไม่จำเป็นต้องเท่ากับศูนย์เสมอไป}$$

2.2 แต่ละส่วนมีคะแนนสอบ (X) เฉลี่ยต่างกัน

2.3 ความแปรปรวนของคะแนนสอบ (X) ต่างกันเล็กน้อย

นักทฤษฎี ที่เสนอเทคนิคสำหรับการประมาณค่าความเชื่อมั่นตามกลุ่มนี้ ได้แก่ ฟลานาแกน (Flanagan) รูลอน (Rulon) กัตแมน (Guttman) เฟลด์ท์ และ เบรนนอน (Feldt & Brennan) และครอนบาค (Cronbach)

3. ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)

เป็นการประมาณค่าความเชื่อมั่นที่ผ่อนปรนเงื่อนไขข้อต่าง ๆ โดยคงไว้เฉพาะเงื่อนไขข้อ 1 ที่ว่าแต่ละส่วนของแบบทดสอบเนื้อหาต้องเป็นเอกพันธ์ หรือคุณลักษณะเดียวกัน (บุญเชิด ภิญญอนันตพงษ์, 2536 : 12 ; อ้างอิงจาก Kristof, 1974 : 492)

ลักษณะสำคัญของความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ คือ การแบ่งแบบทดสอบออกเป็น ส่วน ๆ ที่มีขนาดความยาวไม่เท่ากัน หรือมีขนาดความเท่ากันแต่มีการกระจายของคะแนนในแต่ละส่วนแตกต่างกันมาก การประมาณค่าความเชื่อมั่นโดยวิธีนี้ เนื่องจากในทางปฏิบัติจริงบางครั้งต้องแบ่งส่วนให้เหมาะสมตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีจำนวนข้อไม่เท่ากัน ซึ่งส่งผลต่อเงื่อนไขข้อ 5 และ 6 จึงได้มีการพัฒนาเทคนิคที่มีการผ่อนปรนมากที่สุดโดยคงไว้เฉพาะเงื่อนไขข้อ 1 เท่านั้น

การประมาณค่าความเชื่อมั่นนอกจากจะใช้สถิติขั้นสูงวิเคราะห์ด้วยเครื่องคอมพิวเตอร์แล้ว

นักทฤษฎีการวัดยังได้เสนอเทคนิคที่ใช้เครื่องคำนวณแบบธรรมดาไว้หลายเทคนิค ดังนี้

ฮอรัส (บุญเชิด ภิญญอนันตพงษ์, 2533 : 29 ; อ้างอิงจาก Horst, 1951 : 368 - 371) เสนอสูตรของการแบ่งแบบทดสอบเป็นสองส่วน มีความยาวไม่เท่ากัน ดังนี้

$$r_{tt} = \frac{r\sqrt{r^2 + 4\lambda_1\lambda_2(1-r^2)} - r}{2\lambda_1\lambda_2(1-r^2)}$$

r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
r	แทน	สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบทดสอบส่วนย่อยทั้งสองส่วน
λ_1, λ_2	แทน	สัดส่วนของจำนวนข้อสอบแต่ละส่วนย่อยจากจำนวนข้อสอบทั้งหมด

คริสทอฟ (บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Kristof. 1974 : 491 - 499) เสนอสูตรการประมาณค่าความเชื่อมั่นของแบบทดสอบ เมื่อแบ่งแบบทดสอบเป็นสามส่วน ที่มีขนาดความยาวไม่เท่ากัน แต่ละส่วนมีสหสัมพันธ์ของคะแนนจริงเป็นเส้นตรง ดังนี้

$$r_{tt} = \frac{[s_{x1x2}s_{x1x3} + s_{x1x2}s_{x2x3} + s_{x1x3}s_{x2x3}]^2}{s_x^2(s_{x1x2})(s_{x1x3})(s_{x2x3})}$$

เมื่อ	r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	s_x^2	แทน	ความแปรปรวนของคะแนนทั้งฉบับ
	s_{xixj}	แทน	ความแปรปรวนร่วม

เฟลด์ท์ (บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Feldt. 1975 : 557 - 561) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบที่แบ่งสามส่วนไม่เท่ากันของคริสทอฟ ให้สามารถวิเคราะห์กับแบบทดสอบที่แบ่งสองส่วนไม่เท่ากัน โดยเพิ่มข้อตกลงอีกหนึ่งข้อ คือ ความแปรปรวนของคะแนนความคลาดเคลื่อนต้องเป็นไปตามทฤษฎีการวัดแบบมาตรฐานเดิมและไม่จำเป็นต้องทราบจำนวนข้อของแบบทดสอบ ดังนี้

$$r_{tt} = \frac{[s_{x1x2} / \lambda_1\lambda_2]}{s_x^2}$$

เมื่อ

$$\lambda_1 = \frac{[s_{x1}^2 + s_{x1x2}]}{s_x^2}$$

$$\lambda_2 = \frac{[s_{x2}^2 + s_{x1x2}]}{s_x^2}$$

r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
$\lambda_1 \lambda_2$	แทน	สัดส่วนของจำนวนข้อแต่ละส่วนย่อย
S_x^2	แทน	ความแปรปรวนของคะแนนทั้งฉบับ
S_{x1}^2	แทน	ความแปรปรวนของคะแนนย่อยส่วนที่ 1
S_{x2}^2	แทน	ความแปรปรวนของคะแนนย่อยส่วนที่ 2
S_{xixj}	แทน	ความแปรปรวนร่วม

ราชู (บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Raju. 1977 : 549 - 565) ได้พัฒนาสูตรการประมาณค่าความเชื่อมั่นเมื่อแบ่งแบบทดสอบออกเป็นหลาย ๆ ส่วนแต่ละส่วนมีความยาวไม่เท่ากัน เรียกว่า สัมประสิทธิ์เบต้าเค (Coefficient - β_k) แต่ถ้าแต่ละส่วนมีจำนวนข้อเท่ากันแล้วค่าสัมประสิทธิ์เบต้าเคจะเท่ากับสัมประสิทธิ์แอลฟาของครอนบัค ดังนี้

$$\beta_k = \left[\frac{1}{1 - \sum \lambda_i^2} \right] \left[\frac{s_x^2 - \sum s_i^2}{s_x^2} \right]$$

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

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

กิลเมอร์ และเฟลด์ท (บุญเชิด ภิญโญนนตพงษ์. 2537 : 26 - 30 ; อ้างอิงจาก Gilmer and Feldt. 1983 : 101 - 105) ได้เสนอสูตรการประมาณค่าความเชื่อมั่นโดยพยายามลดจำนวนไม่ทราบค่า $K(K + 1) / 2 + 1$ ให้เหลือเท่ากับความสัมพันธ์ของตัวแปร $2K + 1$ แล้วแก้สมการตัวไม่ทราบค่า $2K + 1$ ซึ่งมีสูตรวิเคราะห์ 2 สูตร คือ r_{F1} และ r_{F2} สามารถใช้คำนวณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งเป็นหลาย ๆ ส่วน ($K > 3$) ดังนี้

$$r_{F1} = \frac{\sigma_T^2}{\sigma_X^2}$$

เมื่อ	r_{F1}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	σ_T^2	แทน	ความแปรปรวนของคะแนนจริงทั้งฉบับ
	σ_X^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ

และ

$$r_{F2} = \frac{[\sum D_g]^2}{[\sum D_g]^2 - \sum D_g^2} \left[\frac{\sigma_X^2 - \sum \sigma_g^2}{\sigma_X^2} \right]$$

D_g หาได้จากสูตร ต่อไปนี้

$$D_1 = \frac{\sum_{g \neq 1} S_{1g} - S_{1L}}{\sum_{g \neq L} S_{Lg} - S_{1L}}$$

...

$$D_K = \frac{\sum_{g \neq K} S_{Kg} - S_{KL}}{\sum_{g \neq L} S_{Lg} - S_{KL}}$$

r_{F2}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
S_{KL}	แทน	ความแปรปรวนร่วมระหว่างส่วนย่อย K และ L
L	แทน	แถวที่ผลรวมของความแปรปรวนร่วมมีค่าสูงสุด
σ_g^2	แทน	ความแปรปรวนของคะแนนสอบแต่ละฉบับ
σ_X^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ
S_{Kg}	แทน	ผลรวมของความแปรปรวนร่วมของแถวที่ K ลบด้วยความแปรปรวนร่วมหนึ่งตัวในแถวที่ K
S_{Lg}	แทน	ผลรวมของความแปรปรวนร่วมของแถวที่ L ลบด้วยความแปรปรวนร่วมหนึ่งตัวในแถวที่ L

เลี้ยว (บุญเชิด วิทยุโณนันทพงษ์ 2537 : 33 - 37 ; อ้างอิงจาก Liou, 1989 : 153 - 163) ได้ทำการวิจัย และเสนอสัมประสิทธิ์ค่าความเชื่อมั่นขึ้นมาใหม่สองสัมประสิทธิ์ คือ r_{L1} และ r_{L2} ซึ่งมีข้อดีที่สามารถคำนวณด้วยเครื่องคำนวณธรรมดา และยังสามารถพิสูจน์ได้ทั้งเชิงทฤษฎีและจำนวนตัวเลขว่ามีค่าเท่ากับสัมประสิทธิ์ r_{LK} ของคริสทอฟ (Kristof) และ r_{LF} และของเฟลด์ท (Feldt) เมื่อ K เท่ากับสามหรือสองตอนตามลำดับ ซึ่งมีสูตรในการคำนวณ ดังนี้

$$r_{L1} = \left[\frac{\sum Q_g}{S_x^2} \right] + \left[1 - \frac{\sum S_g^2}{S_x^2} \right]$$

$$Q_1 = \frac{(\sum S_{g1})^2 - \sum S_{g1}^2}{\sum \sum_{g \neq h} S_{gh}}, g, h \neq 1$$

...

$$Q_K = \frac{(\sum S_{gK})^2 - \sum S_{gK}^2}{\sum \sum_{g \neq h} S_{gh}}, g, h \neq K$$

เมื่อ	r_{L1}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	S_{gh}	แทน	ความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ
	S_g^2	แทน	ความแปรปรวนของคะแนนในแต่ละตอน
	S_x^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ
	Q_g	แทน	ความแปรปรวนส่วนย่อยแต่ละส่วน
	S_{gk}	แทน	ความแปรปรวนร่วมระหว่างส่วนย่อยแต่ละส่วน

สูตร r_{L2} เมื่อแบ่งแบบทดสอบเป็น k ส่วน

$$r_{L2} = \left[\frac{1}{1 - \sum \lambda_g^2} \right] \left[1 - \frac{\sum S_g^2}{S_x^2} \right]$$

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

...

$$\lambda_k = \frac{\sum S_{kg}}{S_x^2}$$

เมื่อ	r_{L2}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	S_g^2	แทน	ความแปรปรวนของคะแนนในแต่ละตอน
	S_x^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ
	λ_g^2	แทน	ผลรวมของความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ

บุญเชิด ภิญญอนันตพงษ์ (2538 : 51) ได้เสนอวิธีการประมาณค่าความเชื่อมั่นโดยพัฒนามาจากสัมประสิทธิ์ r_{L2} ของ Liou และให้ชื่อว่าสัมประสิทธิ์ r_B (Coefficient r_B) มีสูตรดังนี้

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

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

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

เมื่อ	$\sum X^2$	แทน	ผลรวมทั้งหมดของคะแนนแต่ละตัวยกกำลังสอง
	$\sum X$	แทน	ผลรวมทั้งหมดของคะแนน
	N	แทน	จำนวนข้อ

ค่า λ_i หาได้จากสูตรต่อไปนี้

$$\lambda_1 = [\sum B_{1i} - P_1(\bar{X} - 1)] / S_x^2$$

$$\lambda_2 = [\sum B_{2i} - P_2(\bar{X} - 1)] / S_x^2$$

...

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

$\bar{X}$	แทน	ผลรวมของค่าความยากแต่ละข้อ
B_{ji}	แทน	สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใด ๆ ถูกทั้งสองข้อ

การแสดงความมาของสัมประสิทธิ์ r_B

บุญเขต ภิญโญนนตพงษ์ (2542 : 15 - 19) กล่าวถึงสัมประสิทธิ์ r_B ในการประมาณค่าความเชื่อมั่นของแบบทดสอบแบบคะแนนจริงสัมพันธ์ แบบมาตรฐานเดิม ที่ได้พัฒนามาจากสัมประสิทธิ์ r_{L2} ของเลียว (บุญเขต ภิญโญนนตพงษ์, 2542 : 15 - 19 ; อ้างอิงจาก Liou, 1989) ที่มี K ข้อ ให้คะแนนแบบ (0 - 1) หรือตอบถูกให้ 1 ตอบผิด หรือไม่ตอบให้ 0 โดยแสดงความมาของสัมประสิทธิ์ r_B ดังนี้

กำหนดให้แบบทดสอบประกอบด้วยข้อสอบ K ข้อ คะแนนสอบ (Observed scores) คะแนนจริง และคะแนนคลาดเคลื่อนของแต่ละข้อ และของทั้งฉบับแทนด้วยสัญลักษณ์ ดังนี้

คะแนนสอบแต่ละข้อและทั้งฉบับแทนด้วย $X_1, X_2, X_3, \dots, X_K$ และ X

คะแนนจริงแต่ละข้อและทั้งฉบับแทนด้วย $T_1, T_2, T_3, \dots, T_K$ และ T

คะแนนคลาดเคลื่อนแต่ละข้อและทั้งฉบับแทนด้วย $E_1, E_2, E_3, \dots, E_K$ และ E

ดังนั้น (1) $X = X_1 + X_2 + X_3 + \dots + X_K$

(2) $T = T_1 + T_2 + T_3 + \dots + T_K$

(3) $E = E_1 + E_2 + E_3 + \dots + E_K$

จากแบบจำลองคะแนนของทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory Model) ได้คะแนนแต่ละข้อดังนี้

(4) $X_i = T_i + E_i$; $i = 1$ ถึง K

และกำหนดข้อตกลงว่า แบบทดสอบประกอบด้วยข้อสอบที่วัดคุณลักษณะเดียวกัน และเป็นไปตามข้อตกลงแบบคะแนนจริงสัมพันธ์ คือ คะแนนจริงของแต่ละข้อมีความสัมพันธ์เป็นเส้นตรงกับคะแนนจริงทั้งฉบับ

(5) $T_i = \lambda_i T + C_i$; $i = 1$ ถึง K

เมื่อ λ_i เป็นลักษณะประจำข้อข้อที่ i ซึ่งกำหนดให้มีค่าเป็นจำนวนจริงบวก

C_i เป็นค่าคงที่ของข้อสอบ i

โดยกำหนดเงื่อนไขว่า $\sum \lambda_i = 1$ และ $\sum C_i = 0$ นำสมการ (5) แทนลงในสมการ (4) ได้

(6) $X_i = \lambda_i T + C_i + E_i$; $i = 1$ ถึง K

ถ้านำแบบทดสอบไปสอบกับกลุ่มตัวอย่างจำนวนหนึ่ง และหาค่าความแปรปรวนของข้อสอบแต่ละข้อ (σ_i^2) และหาค่าความแปรปรวนของข้อสอบแต่ละคู่ (σ_{ij}) ได้ ดังนี้

(7) $\sigma_i^2 = \lambda_i^2 \sigma_T^2 + \sigma_{E_i}^2$

(8) $\sigma_{ij} = \lambda_i \lambda_j \sigma_T^2$; ($j \neq i$)

ผลรวมของความแปรปรวนของคะแนนสอบทุกข้อ คือ

(9) $\sum \sigma_i^2 = \sum \lambda_i^2 \sigma_T^2 + \sum \sigma_{E_i}^2$

นำ σ_x^2 ไปหารสมการ (9) และนิยามความเชื่อมั่นของแบบทดสอบ $\rho_{xx'} = \sigma_T^2 / \sigma_x^2$

จะได้สูตรความเชื่อมั่นซึ่งราชู (บุญเชิด ภิญโญอนันตพงษ์. 2542 : 15 - 19 ; อ้างอิงจาก Raju. 1977) เป็นผู้เสนอไว้เป็นคนแรก ดังนี้

$$(10) \quad \rho_{xx'} = \frac{1}{1 - \sum \lambda_i^2} \left\{ 1 - \frac{\sum \sigma_i^2}{\sigma_x^2} \right\}$$

เลียว (Liou) ได้อาศัยวิธีการของ เฟลด์ท์ (Feldt) ที่จะประมาณค่า λ_i โดยพิจารณาว่าข้อสอบแต่ละข้อคือแบบทดสอบที่ลดจำนวนข้อลงจนเหลือข้อสอบเพียงข้อเดียว ดังนั้นข้อสอบก็คือแบบทดสอบฟอร์มหนึ่งนั่นเอง ซึ่งตามทฤษฎีการทดสอบแบบมาตรฐานเดิม ความแปรปรวนของคะแนนคลาดเคลื่อนแต่ละข้อและ ทั้งฉบับจะมีความสัมพันธ์กันดังนี้ $\sigma_{Ei}^2 = \lambda_i \sigma_E^2$ และถ้านำคะแนนสอบของแต่ละข้อมาสร้างเป็นเมตริกซ์ความแปรปรวน ความแปรปรวนรวมแล้ว จะสามารถหาค่า λ_i ของแต่ละข้อจากแต่ละแถว (Row) ของเมตริกซ์ดังกล่าวได้โดยนำค่า σ_x^2 ไปหารผลรวมของความแปรปรวน ความแปรปรวนรวมในแต่ละแถว ดังนี้

$$(11) \quad \begin{cases} \lambda_1 = \sum \sigma_{1i}^2 / \sigma_x^2 \\ \lambda_2 = \sum \sigma_{2i}^2 / \sigma_x^2 \\ \dots \\ \lambda_k = \sum \sigma_{ki}^2 / \sigma_x^2 \end{cases}$$

จากสมการ (11) นำมาเขียนใหม่ โดยแยกความแปรปรวนรวมของคะแนนตัวมันเองออกมาซึ่งก็คือ ค่าความแปรปรวนของข้อนั้น เช่น

$$(12) \quad \lambda = \frac{\sigma_1^2 + \sum \sigma_{1i}^2}{\sigma_x^2} (i \neq 1)$$

ถ้าสามารถหาค่าความแปรปรวนของคะแนนข้อ 1 กับข้ออื่น ๆ ทั้งหมดได้ ก็สามารถหาค่า λ_i ได้ การหาค่าความแปรปรวนรวมของข้อสอบแต่ละข้อจากสูตรโดยตรงมีความยุ่งยากและใช้เวลามาก วิธีที่ง่าย และรวดเร็วกว่า คือ อาศัยค่านิยามของสัมประสิทธิ์สหสัมพันธ์

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad \text{ดังนั้นความแปรปรวนรวมของข้อสอบข้อ 1 และ ข้อ 2 ได้จากการนำค่าส่วนเบี่ยงเบน}$$

มาตรฐานคูณค่าสัมประสิทธิ์สหสัมพันธ์ คือ $\sigma_{12} = \sigma_1 \sigma_2 \rho_{12}$ แต่เนื่องจากข้อสอบแต่ละข้อให้คะแนนในระบบ (0 - 1) ดังนั้น

(13) $\sigma_1 = \sqrt{P_1(1 - P_1)}$ และ $\sigma_2 = \sqrt{P_2(1 - P_2)}$

เมื่อ P เป็นสัดส่วนของผู้ตอบถูก หรือความยากง่ายของข้อสอบ

สำหรับสัมประสิทธิ์สหสัมพันธ์ จะต้องใช้สูตรกรณีเฉพาะของสัมประสิทธิ์สหสัมพันธ์ของ Pearson คือ สัมประสิทธิ์ ϕ ซึ่งมีสูตร ดังนี้

(14)
$$\phi = \frac{AD - BC}{\sqrt{(A + B)(C + D)(A + C)(B + D)}}$$

เมื่อ A, B, C และ D เป็นสัดส่วนจำนวนผู้ตอบข้อสอบสองข้อถูก หรือ ผิด ดังแสดงในตารางข้อ 2

		ข้อ 2		
		1	0	
ข้อ 1	1	A	B	A + B
	0	C	D	C + D
		A + C	B + D	1.0

จากตารางจะได้ว่า

$A + B = P_1$
 $B + D = P_2$
 $C + D = 1 - P_1$
 $A + C = 1 - P_2$

และจัดค่าต่าง ๆ อยู่ในพจน์ของ P_1 , P_2 และ B จะได้

$A = P_1 - B$
 $D = P_2 - B$
 $C = 1 - P_1 - P_2 + B$

นำค่าดังกล่าวแทนในสมการ (14) ได้

$$\phi_{12} = \frac{B_{12}(1 - P_1 - P_2 + B_{12}) - (P_1 - B_{12})(P_2 - B_{12})}{\sqrt{P_1(1 - P_1)(1 - P_2)P_2}}$$

(15)
$$\phi_{12} = \frac{B_{12} - P_1P_2}{\sqrt{P_1(1 - P_1)}\sqrt{P_2(1 - P_2)}}$$

ดังนั้นจากสมการ (13) และ (15) ได้ความแปรปรวนร่วม ดังนี้

$$\sigma_{12} = \frac{\sqrt{P_1(1-P_1)}\sqrt{P_2(1-P_2)}}{\sqrt{P_1(1-P_1)}\sqrt{P_2(1-P_2)}} \frac{B_{12} - P_1P_2}{\sqrt{P_1(1-P_1)}\sqrt{P_2(1-P_2)}}$$

$$\sigma_{12} = B_{12} - P_1P_2$$

ทำนองเดียวกัน

$$\sigma_{13} = B_{13} - P_1P_3$$

$$\sigma_{14} = B_{14} - P_1P_4$$

หรือเขียนในรูปทั่วไปได้ ดังนี้

$$\sigma_{1i} = B_{1i} - P_1P_i$$

$$\sigma_{2i} = B_{2i} - P_2P_i$$

...

$$\sigma_{Ki} = B_{Ki} - P_KP_i$$

ดังนั้นจากสมการ (11) และ (12) นำมาเขียนใหม่ได้ว่า

$$\lambda_1 = \frac{P_1(1-P_1) + \sum (B_{1i} - P_1P_i)}{\sigma_x^2} \quad ; i \neq 1$$

$$= \frac{P_1(1-P_1) + \sum B_{1i} - P_1 \sum P_i}{\sigma_x^2} \quad ; i \neq 1$$

$$= \frac{P_1(1-P_1 - \sum P_i) + \sum B_{1i}}{\sigma_x^2} \quad ; i \neq 1$$

$$= \frac{P_1(1-\bar{X}) + \sum B_{1i}}{\sigma_x^2} \quad ; i \neq 1$$

$$\lambda_1 = \frac{\sum B_{1i} - P_1(\bar{X} - 1)}{\sigma_x^2} \quad ; i \neq 1$$

ทำนองเดียวกัน

$$\lambda_2 = \frac{\sum B_{2i} - P_2(\bar{X} - 1)}{\sigma_x^2} \quad ; i \neq 2$$

...

$$\lambda_k = \frac{\sum B_{ki} - P_k(\bar{X} - 1)}{\sigma_x^2} \quad ; i \neq K$$

ดังนั้น การประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม (Classical Congeneric) จากแบบทดสอบที่ให้คะแนนแบบ (0 - 1) จำนวน K ข้อ โดยใช้สัมประสิทธิ์ r_B ที่สร้างขึ้นใหม่ จึงมีสูตร ดังนี้

$$(15) \quad r_B = \frac{1}{1 - \sum \lambda_i^2} \left[1 - \frac{\sum P_i(1 - P_i)}{S_x^2} \right]$$

เมื่อ P_i แทน ค่าความยากง่ายของข้อสอบแต่ละข้อ
 S_x^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ

$$\lambda_1 = [\sum B_{1i} - P_1(\bar{X} - 1)] / S_x^2$$

$$\lambda_2 = [\sum B_{2i} - P_2(\bar{X} - 1)] / S_x^2$$

...

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

$$\bar{X} = \sum P_i$$

B_{ji} แทน สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใด ๆ ถูกทั้งคู่

นอกจากนี้ยังมีนักวิจัยอีกกลุ่มหนึ่งเสนอสูตรที่ใช้คำนวณค่าความเชื่อมั่น โดยให้ความสำคัญกับการนำค่าน้ำหนักองค์ประกอบ (Factor Loading) มาใช้ในการคำนวณค่าความเชื่อมั่นของแบบทดสอบ เช่น ไฮเซ และบอร์นสเต็ด (Heise and Lohrnstedt) ได้เสนอสูตร Ω ซึ่งเป็นสัมประสิทธิ์ค่าความเชื่อมั่นของแบบทดสอบเมื่อกำหนดน้ำหนักความสำคัญเท่ากัน ต่อมาอัลเลน (Allen) ได้พัฒนาสูตรของไฮเซ และบอร์นสเต็ดเมื่อพบว่าการกำหนดน้ำหนักความสำคัญไม่เท่ากันจะส่งผลต่อการประมาณค่าความเชื่อมั่น และความคลาดเคลื่อนพร้อมทั้งได้เสนอสูตร Ω_w ซึ่งเป็นสัมประสิทธิ์ค่าความเชื่อมั่นของแบบทดสอบ เมื่อกำหนดน้ำหนักความสำคัญไม่เท่ากันไว้ดังนี้ (บุญเชิด ภิญโญนนตพงษ์. 2538 : 51)

$$\Omega_w = \frac{\sum_{i=1}^n \frac{\lambda_i^2}{1 - \lambda_i^2}}{1 + \sum_{i=1}^n \frac{\lambda_i^2}{1 - \lambda_i^2}}$$

เมื่อ	Ω_w	แทน	สัมประสิทธิ์ค่าความเชื่อมั่น
	λ_i^2	แทน	น้ำหนักองค์ประกอบของข้อคำถามข้อที่ i
	n	แทน	จำนวนข้อคำถาม

ล้วน สายยศ และอังคณา สายยศ (2539 : 209 - 245) กล่าวว่าความเชื่อมั่นเป็นความคงที่ของคะแนนที่ได้จากการสอบนักเรียนคนเดียวกันหลายครั้งในแบบทดสอบชุดเดิม มีค่าอยู่ระหว่าง -1 ถึง +1 โดยจะพิจารณาเฉพาะค่าที่เป็นบวกเท่านั้น ค่าความเชื่อมั่น จะคำนวณในรูปการประมาณค่าที่หาในรูปของค่าสัมประสิทธิ์ของความเชื่อมั่น แบ่งเป็น 2 ประเภท คือ ความเชื่อมั่นของแบบทดสอบอิงกลุ่ม (Reliability of Norm - Referenced Test) กับความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ (Reliability of Criterion Norm - Referenced)

1. ความเชื่อมั่นของแบบทดสอบอิงกลุ่ม มีวิธีคำนวณค่าสัมประสิทธิ์อยู่ 3 วิธี คือ

1.1 ความเชื่อมั่นแบบความคงที่ของคะแนน (Stability Reliability) เป็นการประมาณค่าความเชื่อมั่น โดยพิจารณาจากคะแนนที่ได้จากการทดสอบ 2 ครั้งในเครื่องมือวัดชุดเดียวกัน โดยใช้สูตรคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Product -Moment Coefficient Correlation) ถ้ามีค่าเหมือนเดิมแสดงว่ามีความคงที่ของคะแนน เป็นวิธีการทดสอบซ้ำ (Test Retest Method) การหาความเชื่อมั่นวิธีนี้มีจุดอ่อนตรงช่วงเวลาการสอบอาจทำให้การประมาณค่าความเชื่อมั่นของแบบทดสอบต่ำเกินไปหรือสูงเกินไป ซึ่งผลของการเปลี่ยนแปลงค่าความเชื่อมั่นนี้เรียกว่าเกิด Carry - Over Effect

1.2 ความเชื่อมั่นโดยใช้แบบทดสอบที่เหมือนกันสองฉบับ (Equivalent - Form) เป็นการประมาณค่าความเชื่อมั่นโดยการใช้แบบทดสอบที่มีลักษณะวัดสิ่งเดียวกัน หรือคู่ขนาน (Parallel Forms) ไปสอบกับนักเรียนกลุ่มเดียวกัน จากนั้นจึงนำคะแนนที่ได้ไปคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ โดยสูตรของเพียร์สันการหาความเชื่อมั่นแบบนี้มีปัญหาตรงการสร้างเครื่องมือวัดให้คู่ขนานกันเพราะเป็นเรื่องค่อนข้างยากที่จะสร้างเครื่องมือวัดให้คู่ขนานกันได้

1.3 ความเชื่อมั่นโดยใช้ความสอดคล้องภายใน (Internal Consistency Reliability) เป็นการหาความเชื่อมั่นที่ใช้แบบทดสอบฉบับเดียว ทำการทดสอบเพียงครั้งเดียว มีวิธีการประมาณค่าความเชื่อมั่นดังนี้

1.3.1 วิธีแบ่งครึ่งแบบทดสอบ (Split - half Method) โดยวิเคราะห์ค่าความยากของแบบทดสอบเป็นรายข้อ แล้วเรียงข้อสอบจากข้อง่ายไปข้อยากนำไปทดสอบกับนักเรียนแล้วแบ่งแบบทดสอบเป็นสองส่วนโดยแบ่งให้แต่ละส่วนมีลักษณะเป็นคู่ขนานกันได้ฉบับข้อคู่กับฉบับข้อคี่หาค่าสหสัมพันธ์ระหว่าง

คะแนนโดยสูตรของเพียร์สันได้ค่าความเชื่อมั่นเพียงครั้งฉบับ แล้วหาความเชื่อมั่นทั้งฉบับ โดยสูตรปรับขยายของสเปียร์แมน - บราวน์ (Spearman - Brown Formula)

1.3.2 วิธีคูเดอร์ - ริชาร์ดสัน (Kuder - Richardson Procedure) เครื่องมือที่จะหาความเชื่อมั่นวิธีนี้ต้องมีลักษณะที่วัดองค์ประกอบร่วมกัน และคะแนนแต่ละข้อต้องอยู่ในลักษณะที่ทำถูกต้องได้

1 คะแนน ทำผิดได้ 0 คะแนน วิธีนี้มีสูตรที่ใช้หาความเชื่อมั่นอยู่ 2 สูตร คือ สูตร KR - 20 กับสูตร KR - 21

1.3.3 วิธีของครอนบัก (Cronbach Alpha Procedure) ใช้หาค่าความเชื่อมั่นกับเครื่องมือที่ตรวจให้คะแนนเป็นลักษณะใดก็ได้

1.3.4 วิธีของฮอยท์ (Hoyt 's ANOVA Procedure) การหาความเชื่อมั่นโดยวิธีนี้เหมาะสำหรับเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลประเภทตรวจให้คะแนนต่าง ๆ กัน ในแต่ละข้อและใช้หลักสถิติของการวิเคราะห์ความแปรปรวน

1.4 ความเชื่อมั่นที่มีผู้ให้คะแนนมากกว่า 1 คน เป็นการหาความเชื่อมั่นที่ใช้แบบทดสอบฉบับเดียวทำการทดสอบเพียงครั้งเดียว และมีผู้ตรวจให้คะแนนมากกว่า 1 คน ซึ่งหาในรูปของสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient : ρ^2) ที่ครอนบัก เกลเซอร์และราชารัตนัม (Cronbach, Gleser and Rajaratnam, 1963) เบรนนัน (Brennan, 1983) ได้เสนอสูตรไว้

1.5 ความเชื่อมั่นของคะแนนผลต่างของแบบทดสอบคะแนนผลต่าง

(Difference Score : $D = x - y$) มีความหมายเป็น 3 ลักษณะ คือ

1.5.1 คะแนนผลต่างระหว่างนักเรียน 2 คนที่สอบแบบทดสอบฉบับเดียวกัน

1.5.2 คะแนนผลต่างระหว่างคะแนนของแบบทดสอบ 2 ฉบับที่สอบกับผู้สอบคนเดียวกัน

1.5.3 คะแนนผลต่างของการสอบ 2 ครั้ง ในแบบทดสอบฉบับเดียวกันของผู้สอบ

คนเดียวกัน (Pretest - Posttest)

1.6 ความเชื่อมั่นจากการแบ่งส่วนย่อยของแบบทดสอบ ในบางครั้งแบบทดสอบที่สร้างจะแบ่งออกเป็นส่วย่อย ๆ แต่ละส่วนย่อย ๆ (Parts) นั้น มีจำนวนข้อสอบไม่เท่ากันหรือมีจำนวนข้อสอบเท่ากัน แต่มีการกระจายคะแนนไม่เท่ากันและส่วนย่อย ๆ นั้นวัดเนื้อหาเดียวกัน การหาความเชื่อมั่นของแบบทดสอบที่เหมาะสม ควรหาในรูปของคอนเจนเนอริก (Congeneric) มีข้อตกลงว่าส่วนย่อยแต่ละส่วนนั้นมีคะแนนจริง (True Score) สัมพันธ์กันในเชิงเส้นตรงการหาความเชื่อมั่นแบบนี้ เป็นการหาความสอดคล้องภายใน โดยเอาคะแนนสอบเพียงครั้งเดียวของแบบทดสอบมาคำนวณ เมื่อแบบทดสอบแบ่งส่วนย่อยเป็นสองส่วน

คำนวณโดยใช้สูตรของราจู (Raju, 1970.) แบบทดสอบแบ่งส่วนย่อยเป็นสามส่วน คำนวณโดยใช้สูตรของคริสทอฟ (Kristof, 1971.) แบบทดสอบแบ่งส่วนย่อยเป็นหลายคำนวณโดยใช้สูตรของราจู (Raju, 1977.) และเฟลด์ท์ (Feldt, 1975.)

อิทธิพลที่มีต่อค่าความเชื่อมั่นของแบบทดสอบที่มาจาก การสอบครั้งเดียว

สูตรที่คำนวณหาความเชื่อมั่นของแบบทดสอบที่มาจาก การสอบเพียงครั้งเดียวหรือคำนวณค่าความสอดคล้องภายในนั้นจำนวนข้อมีผลต่อค่าความเชื่อมั่น เพราะจำนวนข้อสอบมากข้อจะทำให้ค่าความแปรปรวนสูง การคำนวณค่าความเชื่อมั่นจะสูงด้วย และทำนองเดียวกันจำนวนข้อสอบน้อยจะทำให้ความ

แปรปรวนต่ำ การหาค่าความเชื่อมั่นจะมีค่าต่ำด้วย ดังนั้นถ้าต้องการให้แบบทดสอบมีความเชื่อมั่นสูง ๆ จึงควรออกข้อสอบมาก ๆ ข้อ

2. ความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ หมายถึง ผลของคะแนนสอบที่สอบได้มีความคงที่ในการเป็นผู้รอบรู้ในเรื่องที่สอบ มีวิธีหาหลายวิธี ดังนี้

2.1 ความเชื่อมั่นแบบหาความคงที่ของความรอบรู้ (Stability Reliability) เป็นการคำนวณหาความเชื่อมั่นโดยการนำแบบทดสอบอิงเกณฑ์มาสอบซ้ำ 2 ครั้ง แล้วนำคะแนนที่ได้มาหาความคงที่ของการรอบรู้จากการกำหนดคะแนนจุดตัดที่เหมาะสม คำนวณโดยสูตรของ ชรอกและคอสแครลลี

2.2 ความเชื่อมั่นแบบความสอดคล้องในการตัดสินใจ (Decision Consistency Reliability) เป็นการหาความสอดคล้องระหว่างการทดสอบ 2 ครั้งจากแบบทดสอบฉบับเดียว หรือแบบทดสอบที่คู่ขนานกัน 2 ฉบับ

2.3 การหาค่าความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ จากการทดสอบเพียงครั้งเดียว มีหลายวิธี คือ

2.3.1 ความเชื่อมั่นจากสูตรของลิวิสตัน

2.3.2 ความเชื่อมั่นจากสูตรไบโนเมียล (Binomial Formula)

2.3.3 ความเชื่อมั่นโดยใช้วิธีวิเคราะห์ความแปรปรวน

2.3.4 ความเชื่อมั่นโดยวิธีแบ่งครึ่งแบบทดสอบ

2.3.5 ความเชื่อมั่นโดยสูตรของแฮร์ริส

2.4 ความเชื่อมั่นที่คำนึงจุดประสงค์เชิงพฤติกรรมของแบบทดสอบแบบอิงเกณฑ์

การหาค่าความเชื่อมั่นของแบบทดสอบแบบอิงกลุ่มและอิงเกณฑ์มีหลายวิธี ขึ้นอยู่กับจุดประสงค์ในการสร้างแบบทดสอบ ชนิดของแบบทดสอบ ค่าความเชื่อมั่นของแบบทดสอบที่มีผลการวัดที่คงที่แน่นอน เชื่อถือได้ควรมีค่าสูงกว่า 0.70 จากสูตรการหาค่าความเชื่อมั่นดังกล่าว ผู้วิจัยเลือกใช้สูตรสัมประสิทธิ์ r_B ที่เสนอโดยบุญเชิด ภิญโญอนันตพงษ์ (บุญเชิด ภิญโญอนันตพงษ์, 2538 : 51)

6. เอกสารที่เกี่ยวข้องกับการสอนวิทยาศาสตร์ในโรงเรียนประถมศึกษา

6.1 จุดประสงค์ของการสอนวิทยาศาสตร์ในระดับประถมศึกษา

วิทยาศาสตร์ในชั้นประถมศึกษาได้แยกออกเป็นวิชาวิทยาศาสตร์โดยเฉพาะเหมือนชั้นมัธยมศึกษา แต่จะแทรกไว้ในกลุ่มสร้างเสริมประสบการณ์ชีวิตรวมกับเนื้อหาอื่นๆ ในลักษณะบูรณาการ โดยผสมผสานระหว่างวิทยาศาสตร์ สุขศึกษา และสังคมศึกษา จากหลักสูตรประถมศึกษา พ.ศ. 2521 (ฉบับปรับปรุง 2533) ได้ระบุจุดประสงค์ทั่วไปของกลุ่มสร้างเสริมประสบการณ์ชีวิต ดังนี้

เพื่อให้ผู้เรียนได้เรียนรู้เกี่ยวกับมนุษย์และสิ่งแวดล้อมในด้านอนามัย ประชากร การเมือง การปกครอง ศาสนา วัฒนธรรม วิทยาศาสตร์ และเทคโนโลยี สิ่งแวดล้อมทางธรรมชาติ โดยมุ่งให้ผู้เรียนได้เรียนรู้ถึงสภาพ ปัญหา กระบวนการแก้ปัญหา และสามารถนำประสบการณ์เหล่านี้ไปใช้ประโยชน์ต่อการดำรงชีวิต จึงต้องปลูกฝังให้มีคุณลักษณะดังนี้

1. ให้ความสนใจพื้นฐานและปฏิบัติตนให้ถูกต้อง เกี่ยวกับสุขภาพอนามัย ทั้งทางกายและทางจิตใจ ทั้งส่วนบุคคลและส่วนรวม
2. ให้ความรู้และทักษะพื้นฐานเกี่ยวกับสังคมและธรรมชาติ มีนิสัยใฝ่หาความรู้อยู่เสมอในการดำรงชีวิตที่ดี
3. ให้สามารถปรับตัวให้เข้ากับสภาวะแวดล้อมที่กำลังเปลี่ยนแปลง
4. มีทักษะกระบวนการทางวิทยาศาสตร์และสามารถนำความรู้ทางวิทยาศาสตร์และเทคโนโลยีมาใช้ให้เป็นประโยชน์ในชีวิตประจำวันได้
5. มีความรู้ ความเข้าใจเกี่ยวกับความสัมพันธ์ระหว่างมนุษย์กับสิ่งแวดล้อมทั้งที่เป็นธรรมชาติ เทคโนโลยี และทางสังคม
6. มีความเข้าใจ เลื่อมใสในการปกครองระบอบประชาธิปไตยอันมีพระมหากษัตริย์เป็นประมุข
7. เข้าใจหลักของการอยู่ร่วมกันในสังคม โดยตระหนักในหน้าที่ความรับผิดชอบ ปฏิบัติตามขอบเขตแห่งเสรีภาพของตนเองและผู้อื่น
8. มีความภาคภูมิใจในความเป็นไทยและความเป็นเอกราชของชาติ (กระทรวงศึกษาธิการ. 2532 : 29)

6.2 จุดประสงค์การเรียนรู้ กลุ่มสร้างเสริมประสบการณ์ชีวิตใน ป.02 ชั้น ป.6 ตรงตามคู่มือการประเมินผลการเรียน หลักสูตรประถมศึกษา พ.ศ. 2521 (ฉบับปรับปรุง พ.ศ. 2533) กระทรวงศึกษาธิการ

กระบวนการแก้ปัญหา

- * 1. วิเคราะห์ปัญหาสาเหตุของปัญหาและผลกระทบที่จะเกิดจากปัญหา
- * 2. เสนอแนวทางแก้ปัญหาตามข้อมูลและสภาพของปัญหา
- * 3. เลือกแนวทางและวิธีการปฏิบัติในการแก้ปัญหาที่เหมาะสมและมีประสิทธิภาพสูงสุด
- * 4. แก้ไขปัญหาอย่างเป็นระบบ

- * 5. ปรับปรุงการแก้ปัญหาให้เกิดผลทางบวกอย่างมีประสิทธิภาพ
ความรู้ความสามารถและการปฏิบัติในชีวิตประจำวัน
- * 6. นำความรู้และวิธีการทางวิทยาศาสตร์ไปใช้ในชีวิตประจำวัน
- * 7. เห็นความสำคัญของวิทยาศาสตร์
 - 8. ยอมรับและเห็นความจำเป็นในการเสริมสร้างสุขภาพกายและจิต
 - 9. ปฏิบัติตนในการดูแลสุขภาพอย่างสม่ำเสมอทุกวัน
- * 10. วางแนวทางในการทำงานกลุ่ม
- * 11. นำกระบวนการทำงานกลุ่มไปใช้ในการทำงานกลุ่มเป็นประจำ
- * 12. มีเจตคติที่ดีต่องานอาชีพ และปรับปรุงงานอยู่เสมอ
- * 13. วิเคราะห์ผลดีของการปฏิบัติเกี่ยวกับการอนุรักษ์ บำรุงรักษา ศิลปวัฒนธรรม สิ่งแวดล้อม

และทรัพยากรธรรมชาติ

- * 14. ปฏิบัติและแนะนำชักชวนผู้อื่นให้ปฏิบัติในการอนุรักษ์และบำรุงรักษาศิลปวัฒนธรรม สิ่งแวดล้อม และทรัพยากรธรรมชาติ
- * 15. เสนอแนวปฏิบัติ ที่เหมาะสมในการใช้ประโยชน์จากความรู้ ข่าวสาร และข้อมูล
- * 16. นำความรู้ ข่าวสาร ข้อมูล ไปใช้ประโยชน์ในการทำงานและดำเนินชีวิตประจำวันอย่างเหมาะสม

- 17. อธิบายแนวปฏิบัติที่เหมาะสมในการอยู่ร่วมกันในครอบครัวและสังคม
- 18. ปฏิบัติตนตามบทบาทหน้าที่ที่มีต่อครอบครัว ชุมชน ประเทศชาติ ศาสนา และพระมหากษัตริย์ ถูกต้องเหมาะสมและปรับปรุงให้ดีขึ้นอยู่เสมอเป็นประจำ
- 19. เสนอแนวปฏิบัติที่สอดคล้องกับหลักธรรมทางศาสนา
- 20. ปฏิบัติตนและแนะนำชักชวนผู้อื่นให้ปฏิบัติตามหลักธรรมทางศาสนาอยู่เสมอ
- * 21. ปฏิบัติตนให้เหมาะสมและสอดคล้องกับสภาพการเปลี่ยนแปลงของสังคมและสิ่งแวดล้อม

หมายเหตุ * แสดงให้เห็นว่า จากจุดประสงค์การเรียนรู้ กลุ่มสร้างเสริมประสบการณ์ชีวิตในปี. 02 ชั้นป.6 จะตรงกับส่วนเนื้อหาที่เกี่ยวกับวิทยาศาสตร์

เนื้อหาในกลุ่มสร้างเสริมประสบการณ์ชีวิตได้จัดไว้เป็นหน่วย เริ่มจากเรื่องที่ใกล้ตัวนักเรียนและขยายวงกว้างออกไปสู่ชุมชน ชาติ ประเทศเพื่อนบ้าน โลก จักรวาล โดยกำหนดเนื้อหาออกเป็น 12 หน่วย และแบ่งระยะเวลาในการเรียนออกเป็น 3 ช่วง ดังแสดงในตาราง 8 (สมน อมรวิวัฒน์. 2526 :

ตาราง 8 แสดงหน่วยการเรียนรู้และหน่วยการเรียนรู้ย่อย กลุ่มสร้างเสริมประสบการณ์ชีวิต
จำแนกเป็นระดับชั้น

หน่วยที่	ชื่อหน่วย	ช่วงระดับชั้น		
		ป.1 - 2	ป.3 - 4	ป.5 - 6
1	สิ่งมีชีวิต	✓	✓	✓
2	ชีวิตในบ้าน	✓	✓	✓
3	สิ่งที่อยู่รอบตัวเรา	✓	✓	✓
4	ชาติไทย	✓	✓	✓
5	ข่าว เหตุการณ์ และวันสำคัญ	✓	✓	✓
6	การทำมาหากิน		✓	✓
7	พลังงานและสารเคมี		✓	✓
8	จักรวาลและอวกาศ		✓	✓
9	ประเทศเพื่อนบ้าน			✓
10	การสื่อสารและคมนาคม			✓
11	ประชากรศึกษา			✓
12	การเมืองและการปกครอง			✓

7. งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนน

7.1 งานวิจัยในต่างประเทศ

มีผู้ศึกษาค้นคว้าเกี่ยวกับการตรวจให้คะแนนแบบต่าง ๆ ไว้ ดังนี้

คูลมบ์ส และโวมเมอร์ (Coombs, Milholland and Womer. 1956 : 13 - 37) ได้ศึกษาเปรียบเทียบวิธีที่เสนอแนะกับวิธีธรรมดาโดยใช้แบบทดสอบ 3 ฉบับ ๆ ละ 40 ข้อ ได้แก่แบบทดสอบเกี่ยวกับคำศัพท์ ความรู้ของคนขับรถ และแบบทดสอบการใช้สายตาอ่านมิติสัมพันธ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นปีที่ 3 และ 4 ของโรงเรียนมัธยมแจคสัน (Jackson High School) รัฐมิชิแกน จำนวน 855 คน แต่ละกลุ่มมีเงื่อนไขการตอบ และการตรวจให้คะแนนที่แตกต่างกัน คือวิธีของคูลมบ์ส วิธี 0 - 1 และวิธีทั้งสองร่วมกัน ผลการศึกษาพบว่าค่าความเชื่อมั่นจากการตรวจให้คะแนนวิธีของคูลมบ์สสูงกว่าแบบ 0 - 1 ความเที่ยงตรงไม่ต่างกัน วิธี การตรวจให้คะแนนของคูลมบ์สป้องกันการเดาได้ดีกว่า และทัศนคติของผู้สอบชอบวิธีการตรวจของคูลมบ์ส มากกว่า วิธีที่คูลมบ์สเสนอมียอดดี คือ ทำให้พิสัยของคะแนนกว้างขึ้น สามารถวัดความรู้ของนักเรียนได้ละเอียดกว่าวิธี 0 - 1 และยังทราบความรู้ที่ผิดของนักเรียนได้อีกด้วย แต่ยังมีข้อบกพร่องของโอกาสของการได้คะแนนจากการรู้ตัวเลือกที่ผิดและถูกต่างกัน

เดวิส และฟิฟเฟอร์ (Davis and Fiffer. 1959 : 159 - 169) ได้ศึกษาเกี่ยวกับค่าความเชื่อมั่น ค่าความเที่ยงตรงของแบบทดสอบวัดความถนัด และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยกำหนดน้ำหนักของคะแนนแต่ละตัวเลือก กลุ่มตัวอย่างที่ศึกษาเป็นนักเรียนการบินที่ Lackland Air Force Base จำนวน 65 คน ผลการศึกษาพบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนด้วยวิธีให้น้ำหนักคะแนนทุกตัวเลือกสูงกว่าวิธีตรวจให้คะแนนแบบ 0 - 1 และมีความเที่ยงตรงสูงกว่า แต่ข้อจำกัดของการตรวจวิธีนี้ คือ การกำหนดค่าน้ำหนักคะแนนให้แก่ตัวเลือก เพราะข้อสอบที่จะนำมากำหนดให้ค่าน้ำหนักคะแนนของตัวเลือกแตกต่างกันนั้นต้องเป็นข้อสอบที่สร้างโดยผู้ชำนาญและมีความสามารถยอดเยี่ยมในการสร้างข้อสอบ ส่วนการกำหนดค่าน้ำหนักนั้นต้องวิเคราะห์มาจากผลการตอบของนักเรียนมิใช่ผู้ตรวจให้คะแนนเป็นผู้กำหนดขึ้นเอง ซึ่งเป็นข้อจำกัดในทางปฏิบัติ

ฟิฟ และ บรัลซา (Pugh and Brunza. 1975 : 75 - 78) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจให้คะแนน 2 วิธี คือ วิธีธรรมดา (0 - 1) และวิธีให้ผู้ตอบบอกระดับความมั่นใจในการตอบข้อสอบข้อนั้น ซึ่งแบ่งออกเป็น 3 กลุ่มๆ ละ 28 คน เครื่องมือที่ใช้คือแบบทดสอบเลือกตอบวัดสติปัญญา (Intelligence Scale ; L.E.R) จำนวน 48 ข้อ แบ่งเป็น 2 ส่วน ๆ ละ 24 ข้อ โดยศึกษากับนักศึกษามหาวิทยาลัย อินเดียนา (Indiana University) จำนวน 84 คน นำมาทดสอบกลุ่มตัวอย่างทั้ง 3 กลุ่มคือ กลุ่มแรกใช้แบบทดสอบชุด A ตรวจแบบธรรมดาทั้งสองส่วน กลุ่มที่สองใช้แบบทดสอบชุด B ตรวจแบบธรรมดาในส่วนที่หนึ่ง และตรวจโดยให้ผู้ตอบบอกความมั่นใจในส่วนที่สอง สำหรับกลุ่มที่สามใช้แบบทดสอบชุด C ตรวจโดยให้ผู้ตอบบอกความมั่นใจทั้งสองส่วน ผลการศึกษาพบว่าค่าความเชื่อมั่นของแบบทดสอบซึ่งหาโดยวิธีวิเคราะห์ความแปรปรวนของแบบทดสอบที่มีวิธีการตรวจโดยให้ผู้ตอบบอกความมั่นใจมีค่าสูงกว่าการตรวจแบบธรรมดา คือ มีค่าเท่ากับ .85 และ .57 ตามลำดับ

อาบู ซาอ์ฟ (Abu - Sayf . 1975 : 160 - 162) ได้ศึกษาวิธีการตอบโดยให้บอกระดับความมั่นใจ โดยทำการศึกษากับนักเรียนระดับมัธยมศึกษา จำนวน 152 คน ตอบแบบทดสอบชนิดเลือกตอบ ซึ่งมี 4 ตัวเลือก วิชาภาษาอังกฤษ จำนวน 60 ข้อ โดยให้นักเรียนเลือกตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียว และบอกระดับความมั่นใจในการตอบว่าตอบด้วยความมั่นใจ ไม่ค่อยมั่นใจ หรือ เดา แล้วเปรียบเทียบกับเปอร์เซ็นต์ในการตอบถูก หรือ ผิด ผลปรากฏว่าจากคำตอบทั้งหมด การตอบโดยการเดามีโอกาสตอบถูก

อาบู ซาอ์ฟ และไดมอนด์ (Abu - Sayf and Diamond. 1976 : 62 - 65) ได้ศึกษาเกี่ยวกับค่าความเชื่อมั่น และค่าความเที่ยงตรงของแบบทดสอบที่มีวิธีการตรวจให้คะแนน 3 วิธี คือ ให้คะแนนเฉพาะข้อที่ถูกและตอบอย่างมั่นใจ ให้คะแนนข้อที่ถูกโดยไม่ได้เดา และให้คะแนนทุกข้อที่ตอบถูกโดยไม่คำนึงถึงว่ามีความมั่นใจในการตอบถูกหรือไม่เพียงใดโดยใช้สูตรค่าความเชื่อมั่นครอนบัคแอลฟา (Cronbach's Alpha) ผลการศึกษาพบว่า การให้คะแนนเฉพาะข้อที่ถูกและมีความมั่นใจในการตอบส่วนการให้คะแนนทุกข้อที่ตอบถูกโดยไม่ได้เดา มีค่าความเชื่อมั่นในการตอบสูงสุดเท่ากับ .898 ส่วนการให้คะแนนเฉพาะข้อที่ตอบถูกโดยไม่คำนึงถึงความมั่นใจมีความเชื่อมั่น .871 และ .774 ตามลำดับ สำหรับค่าความเที่ยงตรงวิธีให้คะแนนเฉพาะข้อที่ถูกและมั่นใจ ให้คะแนนข้อที่ถูกโดยไม่ได้เดา และวิธีให้คะแนนทุกข้อที่ตอบถูกโดยไม่คำนึงถึงความมั่นใจในการตอบมีค่าความเที่ยงตรง .260, .236 และ .219 ตามลำดับ

กิบบอนส์ และคณะ (Gibbons and others. 1979 : 261) ได้เสนอแนะวิธีตอบและวิธีการตรวจให้คะแนนในแต่ละข้อ โดยให้ผู้ตอบเลือกที่ถูกได้ตั้งแต่ 1 ถึง K - 1 ตัวความมั่นใจ เมื่อ K คือ จำนวนตัวเลือกทั้งหมดในแต่ละข้อ ถ้ามั่นใจว่าตัวเลือกใดถูกนั้นก็ให้ตอบเพียงตัวเลือกเดียวถ้าไม่มั่นใจหลายตัวเลือก

ให้ตอบทุกตัวเลือกที่ไม่มั่นใจ ส่วนวิธีการตรวจให้คะแนนยึดถือหลักว่า คะแนนคาดหวัง (Expected Score) ที่ได้จากการเดาสุ่มจะต้องเป็นศูนย์ในกรณีที่ข้อสอบมี 5 ตัวเลือก และ ผู้ตอบตอบอย่างสุ่ม 2 ตัวเลือก ความน่าจะเป็นที่เขาจะตอบถูกเป็น $2/5$ และ ความน่าจะเป็นที่เขาจะตอบผิดเป็น $3/5$ ถ้าต้องการให้คะแนนคาดหวังเป็นศูนย์ เมื่อเขาตอบถูก จะได้คะแนน 3 คะแนน แต่ถ้าผิดจะได้คะแนน -2 คะแนน

जारาดัท และซาเวจ (Jaradat and Sawaged. 1986 : 369 - 378) ศึกษาเปรียบเทียบค่าความเชื่อมั่น และค่าความเที่ยงตรงของแบบทดสอบวิชาเคมีที่มีการตรวจให้คะแนน 3 วิธี คือ วิธี 0 - 1 วิธีใช้สูตรแก้การเดา และวิธีของกินบอนส์และคณะที่เรียกว่า เทคนิคการเลือกกลุ่มย่อยกลุ่มตัวอย่างที่ใช้ศึกษาเป็น นักศึกษาระดับ 9 ของโรงเรียนในประเทศจอร์แดน จำนวน 160 คน ผลการศึกษพบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนวิธี 0 - 1 สูงกว่าวิธีการตรวจให้คะแนนของกินบอนส์และคณะอย่างมีนัยสำคัญ แต่สูงกว่าวิธีการตรวจให้คะแนนแบบแก้การเดาอย่างไม่มีนัยสำคัญ และค่าความเที่ยงตรงของวิธีตรวจให้คะแนนของกินบอนส์และคณะสูงกว่าวิธีตรวจให้คะแนนวิธี 0 - 1 และวิธีการแก้การเดาอย่างมีนัยสำคัญ

สมบุรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613- A) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์ (ITBS) Form 8 จำนวน 3 ฉบับ คือ ฉบับที่ 1 มโนทัศน์เกี่ยวกับคณิตศาสตร์ ฉบับที่ 2 การแก้โจทย์ปัญหา และฉบับที่ 3 ความรู้ด้านการคิดคำนวณ โดยศึกษากับนักเรียนเกรด 8 ที่ Urban School District ในอเมริกาตะวันตกเฉียงใต้ (Southwestern United States) จำนวน 700 คน ซึ่งแบ่งเป็น 7 กลุ่ม กลุ่มละ 100 คน ซึ่งมีการจัดเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด และตรวจให้คะแนนโดยใช้กฎชิลลิ่งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด และนับคะแนนตั้งแต่ข้อสอบข้อแรกถึงข้อที่ตอบถูกก่อนข้อที่ตอบผิดตำแหน่งนั้น ๆ ข้อสอบข้อที่ตอบถูกหลังข้อที่ตอบผิดจะไม่นำมานับคะแนนด้วยถือว่าเป็นข้อที่ตอบถูกเพราะการเดา ผลการศึกษพบว่า การตอบผิดตำแหน่งที่สองถึงห้า ให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่งและการไม่กำหนดตำแหน่งการตอบผิด โดยการตอบผิดตำแหน่งที่สองทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงสุดทั้ง 3 ฉบับ

วิลสัน และคณะ (Wilson ,et al. 1988 : 213 - 217) ได้ทำการวิจัยเพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ได้จากการตรวจให้คะแนน โดยใช้กฎชิลลิ่ง ที่กำหนดจุดชิลลิ่งต่างกัน 5 จุด กับวิธีการตรวจโดยไม่กำหนดจุดชิลลิ่งจากแบบทดสอบที่ได้เรียงลำดับข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด ซึ่งคะแนนรวมของแบบทดสอบจะนับเฉพาะข้อที่ตอบถูกก่อนถึงข้อที่ตอบผิดเป็นลำดับที่ต่าง ๆ ข้อที่ตอบถูกหลังจากข้อที่ตอบผิดลำดับที่นั้น ๆ จะไม่นำมานับคะแนนรวม ถือว่าเป็นข้อที่ตอบถูกเพราะเกิดจากการเดาในการวิจัยครั้งนี้ได้ทดสอบกับนักเรียนระดับมัธยมศึกษาตอนต้นจำนวน 700 คน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์จำนวน 44 ข้อ ผลปรากฏว่าการตรวจให้คะแนนแบบทดสอบโดยใช้กฎชิลลิ่ง ที่กำหนดลำดับที่ของการตอบผิดลำดับต่าง ๆ กันมีผลต่อความเชื่อมั่นของแบบทดสอบแตกต่างกันและการกำหนดลำดับที่ของการตอบเป็นสองถึงห้า ให้ค่าความเชื่อมั่นของแบบทดสอบสูงกว่า การกำหนดลำดับที่ของการตอบผิดเป็นหนึ่ง และการไม่กำหนดลำดับที่ของการตอบผิด และยังพบว่าการกำหนดลำดับที่ของการตอบผิดเป็นลำดับที่สองมีค่าความเชื่อมั่นสูงกว่าการกำหนดลำดับที่ของการตอบผิดเป็นลำดับที่ห้า

7.2 งานวิจัยภายในประเทศ

มีผู้ศึกษาค้นคว้าวิจัยเกี่ยวกับวิธีการตรวจให้คะแนนแบบต่าง ๆ ดังนี้

อรรวรรณ ตันต์เจริญรัตน์ (2517 : 1 - 67) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีวิธีการตอบและวิธีการตรวจให้คะแนนแบบ 0 - 1 วิธีของคูมบัส และวิธีของอนันต์ ศรีโสภา เครื่องมือที่ใช้เป็นแบบทดสอบวิชาภาษาไทยชนิด 5 ตัวเลือก จำนวน 3 ฉบับ กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาระดับประกาศนียบัตรการศึกษาตอนต้นวิทยาลัยครูหมู่บ้านจอมบึง จังหวัดราชบุรี จำนวน 214 คน สุ่มเป็น 3 กลุ่ม ผลการศึกษาพบว่าแบบทดสอบที่มีวิธีตอบและวิธีตรวจโดยวิธีใหม่ตามแนวคิดของอนันต์มีค่าความเชื่อมั่นของสูงกว่าวิธีอื่น ๆ

ธีรศักดิ์ อินทรมาตย์ (2520 : 63 - 73) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่น ค่าความเที่ยงตรง และปริมาณการเดาแบบทดสอบวิชาวิทยาศาสตร์ ที่มีการตรวจให้คะแนน 3 วิธี คือ วิธีตอบแบบธรรมดา (0 - 1) วิธีบอกความมั่นใจในการตอบ วิธีตอบทุกตัวเลือก ทดสอบกับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนสมุทรสงคราม จำนวน 216 คน โดยแบ่งเป็น 3 กลุ่มแต่ละกลุ่มได้รับแบบทดสอบที่มีวิธีการตอบ และตรวจให้คะแนนแต่ละวิธีตามลำดับ ซึ่งพบว่าวิธีการตอบโดยบอกความมั่นใจในการตอบ มีความเชื่อมั่นสูงกว่าวิธีตอบแบบธรรมดาและวิธีตอบทุกตัวเลือก ปริมาณการเดาของนักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ในระดับต่ำจะตอบแบบทดสอบโดยการเดามากกว่านักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์สูงกว่าในทุกวิธีการตอบ แบบทดสอบที่มีวิธีตอบทุกตัวเลือกมีค่าความเที่ยงตรงสูงสุด วิธีตอบแบบธรรมดาและแบบทดสอบที่มีการตอบโดยบอกความมั่นใจมีความเที่ยงตรงลดลงตามลำดับ แต่ทั้งสามวิธีมีค่าความเที่ยงตรงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เพ็ญศรี สว่างเนตร (2520 : บทคัดย่อ) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบด้วยเทคนิคการให้คะแนนต่างกัน 4 วิธี คือ วิธีกำหนดคะแนนในข้อที่ตอบถูกข้อละหนึ่งคะแนน (0 - 1) วิธีหักคะแนนข้อที่ตอบผิด วิธีเพิ่มคะแนนข้อที่เว้น และวิธีทดสอบความมั่นใจ กลุ่มตัวอย่างที่ศึกษาเป็นนักศึกษาระดับประกาศนียบัตรวิชาการศึกษาชั้นสูงที่เรียนวิชาภาษาอังกฤษเป็นวิชาเอก หรือวิชาโท จำนวน 218 คน และหาค่าความเชื่อมั่นจากสูตรการวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance) ซึ่งพบว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้เทคนิคให้คะแนนความความมั่นใจ (.61) สูงกว่าแบบ 0 - 1 (.55) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเชื่อมั่นของแบบทดสอบที่ใช้เทคนิคลดคะแนนข้อตอบผิด (.56) มีค่าเพิ่มจากแบบ 0 - 1 เล็กน้อย และแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 จึงสรุปได้ว่าเทคนิคให้คะแนนความมั่นใจที่ใช้กับแบบทดสอบเลือกตอบมีค่าความเชื่อมั่นสูงสุด

เกริกชัย ฮวงเจริญ (2525 : 27 - 34) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นและคะแนนการเดาของการตอบ และตรวจให้คะแนนแบบทดสอบปรนัยเลือกตอบ ที่มีวิธีการตอบ 3 วิธี คือ วิธีใหม่ วิธีธรรมดา และวิธีทดสอบความมั่นใจ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2524 โรงเรียนวัดอ่างทอง จำนวน 114 คน แบ่งเป็น 3 กลุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวิชาวิทยาศาสตร์แบบปรนัยเลือกตอบชนิด 5 ตัวเลือก จำนวน 40 ข้อ ผลการศึกษาพบว่าวิธีตอบ และวิธีตรวจให้คะแนนใหม่ และวิธีทดสอบความมั่นใจ ให้ค่าความเชื่อมั่นสูงกว่าแบบธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และให้ค่าคะแนนการเดาของนักเรียนน้อยกว่าวิธีธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

สำราญ มีแจ้ง (2525 : 25 - 61) ได้ศึกษาผลของคำสั่ง และวิธีการให้คะแนนที่ต่างกันต่อค่าความเชื่อมั่น ค่าความเที่ยงตรง และ ค่าอำนาจจำแนกของแบบทดสอบเลือกตอบ ที่มีตัวเลือกถูกตัวเดียวโดยมีวิธีการให้คะแนน 4 วิธีคือ วิธี 0 - 1 วิธีคูมบัส วิธีของอนันต์ และวิธีของสำราญกับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนสุรศักดิ์มนตรี และโรงเรียนวัดน้อยในที่เรียนวิชาคณิตศาสตร์ ค 011

ปีการศึกษา 2524 จำนวน 240 คนโดยเมื่อผู้สอบทำเครื่องหมายตรงกับตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียวแล้วถูกต้องจะได้ 4 คะแนน ผลการศึกษาพบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและวิธีการตรวจให้คะแนนแบบสำราญสูงกว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0 - 1

($P < .01$) และ ค่าความเชื่อมั่นที่ใช้คำสั่ง และการให้คะแนนตามวิธีของคูมบัสสูงกว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0 - 1 ($P < .05$)

สุพรรณ สุขมลสันต์ (2528 : 47 - 60) ศึกษาเปรียบเทียบการให้น้ำหนักคะแนนที่ต่างกัน 5 วิธี คือ 1) วิธีใช้ค่าเฉลี่ย (Δ) 2) วิธีใช้ค่าเบี่ยงแปร (β_1) 3) วิธีใช้น้ำหนักคะแนนที่เหมาะสม (w) 4) วิธีใช้ค่าผลบวกของความแปรปรวนร่วม (h^2) และ 5) วิธีประเพณีนิยมกับแบบทดสอบการอ่านจับใจความภาษาอังกฤษ 2 ฉบับ คือ ฉบับที่ 1 เป็นฉบับที่ไม่มีชื่อเรื่องและภาพประกอบ ฉบับที่ 2 มีชื่อเรื่องและภาพประกอบ กลุ่มตัวอย่างที่ใช้เป็นนิสิตชั้นปีที่ 1 ที่เรียนภาษาอังกฤษพื้นฐานเฉพาะการอ่าน 2 ของจุฬาลงกรณ์มหาวิทยาลัย จำนวน 1,872 คน ผลการศึกษาพบว่าความเชื่อมั่นของแบบทดสอบอิงกลุ่ม คำลำดับของแบบทดสอบทั้ง 2 ฉบับ ที่ให้คะแนนวิธีต่าง ๆ มีค่าสูงทุกวิธี โดยแบบทดสอบที่ให้คะแนนวิธีใช้น้ำหนักคะแนนที่เหมาะสมมีค่าความเชื่อมั่นสูงสุด

พนิตา สวัสดิมงคล (2530 : 63 - 67) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบที่วัดระดับความรู้ที่ต่างกัน 2 ระดับ คือระดับความรู้ขั้นต่ำและความรู้ขั้นสูง โดยมีวิธีการให้คะแนนวิธีธรรมดา กับวิธีการให้คะแนนความมั่นใจ โดยศึกษากับนักศึกษาพยาบาลชั้นปีที่ 3 ปีการศึกษา 2530 หลักสูตรพยาบาลศาสตรบัณฑิต (วิทยาศาสตร์บัณฑิตสาขาพยาบาล) สังกัดทบวงมหาวิทยาลัย และองค์การกุศลในเขตกรุงเทพมหานคร ในกลุ่มผู้สอบที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลางและต่ำ เครื่องมือที่ใช้คือแบบทดสอบวัดผลสัมฤทธิ์ทางการพยาบาลผู้ป่วยระบบทางเดินหายใจ จำนวน 60 ข้อ เป็นข้อสอบที่วัดระดับความรู้ขั้นต่ำ 30 ข้อ วัดระดับความรู้ขั้นสูง 30 ข้อ ผลการศึกษาพบว่า 1) ค่าความเชื่อมั่นของแบบสอบวัดที่ระดับความรู้ขั้นต่ำ และวิธีให้คะแนนแบบธรรมดาจะหว่างกลุ่มต่ำ กับกลุ่มปานกลางและสูง มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 2) ค่าความเชื่อมั่นของแบบสอบวัดที่วัดระดับความรู้ขั้นสูง และมีวิธีให้คะแนนแบบธรรมดาที่ได้จากทั้ง 3 กลุ่มมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 3) ค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นต่ำและมีวิธีให้คะแนนความมั่นใจที่ได้จาก 3 กลุ่ม มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 4) ค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นสูงและมีวิธีให้คะแนนความมั่นใจระหว่างกลุ่มสูงกับกลุ่มต่ำมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผลจากการเปรียบเทียบค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นต่ำและแบบสอบที่วัดระดับความรู้ขั้นสูงที่ได้จากวิธีให้คะแนนทั้ง 2 วิธี มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทั้งในกลุ่มสูง ปานกลางและต่ำ

เพลินพิศ คุณคำ (2532 : 1 - 79) ศึกษาเปรียบเทียบค่าความเชื่อมั่นและปริมาณการเดาของแบบทดสอบปรนัยชนิดเลือกตอบที่มีการตอบและตรวจให้คะแนนวิธี 0 - 1 วิธีของคูมบัส และวิธีของอนันต์ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนปรังค์ภู่ จังหวัด ศรีสะเกษ ภาคเรียนที่ 2 ปีการศึกษา 2531 จำนวน 30 คน ซึ่งเลือกมาโดยการสุ่มแบบธรรมดา (Simple Random Sampling) เครื่องมือที่ใช้เป็น

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ (ว 204) จำนวน 2 ฉบับ มีการตรวจให้คะแนนโดยวิธี 0 - 1, วิธีของคูมบัส และวิธีของอนันต์ ผลการศึกษาพบว่า การตรวจให้คะแนนวิธีของอนันต์ให้ค่าความเชื่อมั่นสูงกว่าวิธีอื่น ๆ ส่วนค่าความเที่ยงตรงของแบบทดสอบที่ตรวจให้คะแนนทั้ง 3 วิธี พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญวิธีของอนันต์มีการเดาต่ำกว่าวิธี 0 - 1 และวิธีของคูมบัส และวิธีของอนันต์มีเปอร์เซ็นต์คะแนนที่ถูกกว่าวิธีของคูมบัส

อนันต์ ศรีโสภา (2532 : 1 - 54) ได้ทำการทดลองแบบใหม่ของการสอบแบบเลือกตอบโดยสุ่มนักศึกษาจำนวน 144 คน แบ่งกลุ่มทดลองเป็น 3 กลุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์แบบเลือกตอบวิชาวิจัยเบื้องต้นจำนวน 3 ฉบับ กลุ่มตัวอย่างที่ 1 ให้ตอบตามรูปแบบวิธีการตอบแบบใหม่ กลุ่มตัวอย่างที่ 2 ให้ตอบตามรูปแบบวิธีการตรวจของคูมบัส กลุ่มตัวอย่างที่ 3 ให้ตอบตามรูปแบบการตอบ 0 - 1 ผลปรากฏว่า คะแนนการเดาคะแนนข้อถูก คะแนนข้อผิดและคะแนนข้อที่ถูกเพียงบางส่วนของวิธีการตรวจให้คะแนนรูปแบบใหม่ ไม่แตกต่างจากวิธี 0 - 1 อย่างมีนัยสำคัญ และค่าความเชื่อมั่นวิธีใหม่สูงกว่าวิธีของคูมบัส และวิธี 0 - 1 อย่างมีนัยสำคัญ ส่วนค่าความเที่ยงตรงทั้ง 3 วิธีไม่แตกต่างกัน โดยทั่วไปวิธีการตอบและตรวจให้คะแนนแบบใหม่มีประสิทธิภาพในการวัดเหนือกว่าวิธี 0 - 1 ซึ่งวิธีแบบใหม่ช่วยให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้นและลดจำนวนการเดาให้น้อยลงทำให้มีประสิทธิภาพในการวัดสูงและมีคุณค่าแก่การศึกษา

จินตนา ต่อทรัพย์สินชัย (2533 : 46 - 49) ศึกษาความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยใช้กฎซัลลิงที่กำหนดจุดซัลลิงต่างกัน 5 จุดกับการไม่กำหนดจุดซัลลิง เครื่องมือที่ใช้ในการศึกษา คือแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดจำนวน 50 ข้อ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนสังกัดกรมสามัญศึกษาจังหวัดปทุมธานี ปีการศึกษา 2532 ผลการศึกษาพบว่า การตรวจให้คะแนนที่กำหนดจุดซัลลิงจุดที่ 1, 2, 3 และ 4 ให้ค่าความเชื่อมั่นต่ำกว่าการตรวจที่กำหนดจุดซัลลิงจุดที่ 5 และการไม่กำหนดจุดซัลลิง อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนการตรวจให้คะแนนโดยกำหนดจุดซัลลิงที่ 5 ให้ค่าความเชื่อมั่นสูงสุดเมื่อเทียบกับวิธีการตรวจแบบไม่กำหนดจุดซัลลิงสำหรับค่าความเที่ยงตรงพบว่าการตรวจแบบไม่กำหนดจุดซัลลิงให้ค่าความเที่ยงตรงสูงสุด

พินิจ อุไรรักษ์ (2533 : 57) ศึกษาผลของวิธีการให้คะแนนที่มีต่อคะแนนสอบต่อค่าความเชื่อมั่นและค่าความเที่ยงตรงจากการให้น้ำหนักคะแนนรายข้อเท่ากัน วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถาม และวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในเขตจังหวัดบุรีรัมย์ ปีการศึกษา 2532 ผลการศึกษาพบว่าค่าความเชื่อมั่นของแบบสอบจากการให้น้ำหนักคะแนนรายข้อเท่ากัน วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถาม และวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบมีค่า เท่ากับ .705, .781, .886 และ .894 ตามลำดับ โดยเฉลี่ยวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถามและวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบมีค่าความเที่ยงตรงสูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน

อัมพร บั้มบั้ม (2533 : 59 - 63) ศึกษาผลของการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกันที่มีต่อค่าการเดา และลักษณะโค้งอินฟอร์เมชันของแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นจำนวน 40 ข้อ ที่มีการจัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดเพชรบุรี ปีการศึกษา 2533 ผลการศึกษาพบว่า ค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งจุดที่ 2,3,4 และ 5 แต่ละจุด แตกต่างกับค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่ไม่กำหนดจุดซึลลิ่งอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 กล่าวคือ ค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนโดยที่กำหนดจุดซึลลิ่งจุดที่ 2,3,4 และ 5 มีค่าต่ำกว่าค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนโดยไม่กำหนดจุดซึลลิ่ง โดยที่ค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 2 มีค่าต่ำสุด และค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งระหว่างจุดที่ 2,3,4 และ 5 แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ การศึกษาโค้งอินฟอร์เมชันของแบบทดสอบพบว่า ในช่วงระดับความสามารถต่ำโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 2,3,4 และ 5 สูงกว่าโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่ไม่กำหนดจุดซึลลิ่ง โดยที่โค้งอินฟอร์เมชันของแบบทดสอบที่จุดซึลลิ่งจุดที่ 2 สูงสุด และในช่วงระดับความสามารถสูง โค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 5 สูงสุด

เยาวลักษณ์ น้อยนาแสง (2535 : บทคัดย่อ) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่น ค่าความเที่ยงตรง อำนาจจำแนก เปอร์เซนต์การเดา เปอร์เซนต์ความถนัดทั้งหมด เปอร์เซนต์ความถนัดบางส่วน และเปอร์เซนต์ไม่มีความถนัดที่ใช้ในการตอบ และตรวจให้คะแนน 3 วิธี คือ วิธี 0 - 1 วิธีของคัมบัส วิธีของอนันต์ กับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดร้อยเอ็ด ปีการศึกษา 2534 ผลการศึกษาพบว่าแบบทดสอบที่มีตัวเลือกถูกตัวเดียวมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มีตัวเลือกถูกหลายตัว เมื่อตรวจด้วยวิธี 0 - 1 และวิธีของคัมบัสมีความเที่ยงตรงสูงกว่าแบบ 0 - 1 ส่วนแบบทดสอบที่มีตัวเลือกถูกหลายตัว เมื่อตรวจด้วยวิธีของคัมบัสและวิธีอนันต์มีค่าความเชื่อมั่นสูงกว่าวิธี 0 - 1 ส่วนค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียน ที่มีตัวเลือกถูกตัวเดียว และมีตัวเลือกถูกหลายตัวเมื่อตรวจให้คะแนนทั้ง 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติโอกาสของการเดาเพศชายเดาน้อยกว่าเพศหญิง และ แบบทดสอบความถนัดทางการเรียนแบบที่มีตัวเลือกถูกตัวเดียวตรวจด้วยวิธีของอนันต์เปอร์เซนต์ความถนัดบางส่วนสูงกว่าวิธีของคัมบัส แต่เปอร์เซนต์ความไม่ถนัดตรวจด้วยวิธีของอนันต์ต่ำกว่าวิธีของคัมบัส

มานะ คิวเจริญ (2539 : 57 - 62) ศึกษาค่าความเชื่อมั่นและความสอดคล้องระหว่างคะแนนของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยาก และพฤติกรรมการเรียนรู้ ที่มีการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกัน โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดกาญจนบุรี ปีการศึกษา 2538 จำนวน 434 คน ผลการศึกษาพบว่า

1. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยากที่ได้จากการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งตำแหน่งที่ 1 มีค่าความเชื่อมั่นสูงสุดและไม่กำหนดตำแหน่งจุดซึลลิ่งมีค่าความเชื่อมั่นต่ำสุด
2. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงตามพฤติกรรมการเรียนรู้ ที่ได้จากการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งตำแหน่งที่ 1 มีค่าความเชื่อมั่นสูงสุด และไม่กำหนดตำแหน่งจุดซึลลิ่งมีค่าความเชื่อมั่นต่ำสุด

3. ค่าความสอดคล้องระหว่างคะแนนของแบบทดสอบคณิตศาสตร์ที่มีการจัดเรียงตามค่าความยาก กับเรียงตามพฤติกรรมการเรียนรู้ที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งตำแหน่งเดียวกันทุกตำแหน่ง มีค่าความสอดคล้องกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดย แบบทดสอบที่ไม่กำหนดตำแหน่งจุดซึลลิ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสูงสุด และจุดซึลลิ่งตำแหน่งที่หนึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนต่ำสุด

พิมพร ภูมิรา (2544 : บทคัดย่อ) ศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกัน 3 แบบ กลุ่มตัวอย่างที่ศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 จังหวัดกาญจนบุรี จำนวน 790 คน ผลการศึกษาพบว่า

1. การตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่ 2 และตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .93 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

2. การคะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สองและตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .95 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

3. การคะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .93 - .95 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

4. การตรวจให้คะแนนโดยไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นเฉลี่ยต่ำสุดเท่ากับ .78

5. การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

6. การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

7. การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ชลินทร์ ศิลปสมศักดิ์ (2544 : บทคัดย่อ) ได้ศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีการตรวจให้คะแนนตามกฎซึลลิ่ง โดยมีการจัดเรียงข้อสอบและกำหนดจุดซึลลิ่งต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 จังหวัดสมุทรปราการ จำนวน 852 คน ผลการศึกษาพบว่า

1. ค่าความเชื่อมั่นระหว่างแบบทดสอบที่จัดเรียงข้อสอบตามค่าความยากง่าย กับแบบทดสอบที่จัดเรียงข้อสอบตามพฤติกรรมการเรียนรู้ที่จุดซึลลิ่งที่หนึ่งมีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จุดซึลลิ่งที่สี่ และจุดซึลลิ่งที่ห้า ค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ค่าความเชื่อมั่นระหว่างแบบทดสอบเลือกตอบถูกผิดที่จัดเรียงข้อสอบตามค่าความยากง่าย ระหว่างการกำหนดจุดซึลลิ่งที่หนึ่งและจุดซึลลิ่งที่สอง จุดซึลลิ่งที่สาม จุดซึลลิ่งที่สี่ จุดซึลลิ่งที่ห้า และไม่กำหนดจุดซึลลิ่ง ค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สำหรับความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิด ที่จัดเรียงข้อสอบตามพฤติกรรมกรรมการเรียนรู้ระหว่างการกำหนดจุดซึลลิ่งต่างกันมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

8. งานวิจัยที่เกี่ยวข้องกับการจัดเรียงข้อคำถาม

8.1 งานวิจัยในต่างประเทศ

มีผู้ศึกษาวิจัยเกี่ยวกับการจัดเรียงข้อคำถามไว้ ดังนี้

โมเลนคอฟ (Mollenkopf. 1950 : 291) ได้ศึกษาค้นคว้าผลการจัดเรียงข้อคำถามในแบบทดสอบใหม่ว่าจะมีผลทำให้ค่าทางสถิติเปลี่ยนแปลงหรือไม่ และเปรียบเทียบผลที่เกิดจากการสอบแบบไม่จำกัดเวลา (Power Test) กับ การสอบแบบจำกัดเวลา (Speed Test) พบว่าค่าความยากง่ายของข้อสอบ และค่าอำนาจจำแนกของข้อสอบแต่ละข้อเปลี่ยนแปลงไปจากเดิม แสดงว่าการจัดเรียงข้อคำถามในแบบทดสอบมีผลทำให้ค่าทางสถิติของแบบทดสอบเปลี่ยนแปลงไป

แมคนิคอล (Flangher, Melton and Myers. 1968 : 813 ; citing Mac Nical. 1968) ศึกษาการเปลี่ยนวิธีในการเรียงลำดับข้อคำถามในแบบทดสอบ ซึ่งแต่เดิมเรียงจากง่ายไปหายาก เปลี่ยนไปเป็นเรียงแบบอื่น 2 ชนิด คือ เรียงแบบสุ่ม และเรียงจากข้อยากไปหาข้อง่าย ผลการศึกษาพบว่าการเรียงลำดับข้อคำถามจากข้อยากไปหาข้อง่ายทำให้ข้อสอบยากขึ้นกว่าเดิมอย่างมีนัยสำคัญทางสถิติ แต่การเรียงแบบสุ่มไม่มีผลทำให้ค่าความยากง่ายของข้อสอบแตกต่างไปจากเดิม

แฟลงเกอร์ เมลตันและไมเออร์ส (Flangher, Melton and Myers. 1968 : 813 - 824) ศึกษาการจัดเรียงลำดับของข้อสอบต่อผลคะแนนในการสอบ ข้อสอบที่ใช้ในการวิจัย คือ ข้อสอบด้านภาษาจำนวน 40 ข้อ คณิตศาสตร์ 25 ข้อ ซึ่งข้อสอบทั้งสองชนิดเป็นส่วนหนึ่งของ ซี อี อี บี สกอลาสติก แอปติจูดเทสต์ (C E E B Scholastic Aptitude) มีการจัดแบ่งข้อสอบไว้เป็นตอน ๆ ละ 5 ข้อ จัดเรียงเป็น 4 แบบ คือ

1. จัดเรียงข้อสอบแบบมาตรฐานเดิม (Standard Arrangement)

2. จัดเรียงข้อภายในแต่ละตอน ซึ่งมีตอนละ 5 ข้อใหม่ ส่วนการเรียงระหว่างตอนยังเหมือนมาตรฐานเดิม

3. จัดเรียงลำดับข้อสอบแต่ละตอนใหม่ ส่วนข้อภายในแต่ละตอนนั้นเรียงเหมือนมาตรฐานเดิม

4. จัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอน

ผลการวิจัยพบว่า การจัดเรียงแบบมาตรฐานเดิม และการจัดเรียงข้อภายในแต่ละตอนแบบมาตรฐานเดิมไม่มีผลทำให้คะแนนในการสอบเปลี่ยนแปลง ส่วนการจัดเรียงลำดับข้อสอบแต่ละตอนเสียใหม่ ข้อภายในแต่ละตอนนั้นเรียงเหมือนมาตรฐานเดิม และการจัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอนมีผลทำให้คะแนนในการสอบเปลี่ยนแปลงไป

คลอสเนอร์ และเกลล์แมน (Klosner and Gellman 1973 : 413 - 418) ได้ศึกษาการเรียงลำดับข้อคำถามในแบบทดสอบในชั้นเรียน 3 แบบ คือ โดยเรียงตามหัวเรื่อง เรียงตามลำดับจากง่ายไปยาก และเรียงตามหัวเรื่องซึ่งแต่ละหัวเรื่องเรียงข้อคำถามจากง่ายไปหายาก โดยทดลองกับกลุ่มตัวอย่างที่เป็นนักศึกษาของมหาวิทยาลัยฮอฟสตรา ปี ค.ศ. 1969 จำนวน 54 คน แบ่งออกเป็น 2 กลุ่ม คือ กลุ่มที่มีความสามารถสูง กับกลุ่มที่มีความสามารถต่ำ เครื่องมือที่ใช้เป็นแบบทดสอบวิชาวัดผลการศึกษาเบื้องต้นจำนวน 75 ข้อ ผลการศึกษาพบว่า โดยทั่วไปแล้วการเรียงข้อทดสอบวิธีต่างๆ ไม่มีผลต่อคะแนนการสอบของนักเรียนที่มีความสามารถปานกลาง แต่สำหรับในกลุ่มนักเรียนที่มีความสามารถต่ำแล้ว การเรียงข้อคำถามของแบบทดสอบมีผลทำให้คะแนนการสอบเปลี่ยนไปด้วย และค่าความเชื่อมั่นของแบบทดสอบที่เรียงตามหัวเรื่องและเรียงลำดับจากง่ายไปหายากมีค่าเท่าๆกัน ส่วนในการเรียงจากง่ายไปหายาก ในแต่ละหัวเรื่องนั้นมีความเชื่อมั่นต่ำกว่า 2 วิธีแรก

เพลค (Plake. 1980 : 56 - 58) ศึกษาเกี่ยวกับการจัดเรียงข้อคำถาม 3 แบบ คือ เรียงจากง่ายไปหายาก เรียงแบบยากง่ายหมุนเวียน (Spiral Cyclical) และ เรียงแบบสุ่ม โดยศึกษากับกลุ่มตัวอย่างที่เป็นนักเรียนพยาบาลจำนวน 104 คน พบว่าคะแนนการจัดเรียงข้อคำถามทั้ง 3 แบบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

เลนี และบูล (Lane & Bull. 1987 : 865 - 878) ศึกษาผลของความรู้ เพศ ความยากเชิงสถิติและความยากเชิงปัญญาที่มีต่อการทำข้อสอบที่เรียงข้อคำถามแตกต่างกันตามความยากเชิงสถิติและเรียงตามความยากเชิงปัญญา คือ จำแนกตามจุดมุ่งหมายพฤติกรรมการเรียนรู้ของบลูมชั้นความรู้ความจำ ความเข้าใจ และการนำไปใช้ เครื่องมือที่ใช้เป็นแบบทดสอบแบบเลือกตอบ โดยศึกษากับกลุ่มตัวอย่างที่เป็นนักศึกษา 2 กลุ่ม กลุ่มแรกมีจำนวน 155 คน กลุ่มที่สองมีจำนวน 247 คน ทำการศึกษา 2 ครั้ง ครั้งที่ 1 ใช้แบบทดสอบ 5 ฉบับที่เรียงข้อคำถามต่างกัน ผลการศึกษาพบว่า คะแนนรวมที่ได้จากแบบทดสอบของเพศชายกับเพศหญิงแตกต่างกัน และข้อสอบแต่ละฉบับจัดทำเป็น 2 ชุด ชุดแรกจะมีข้อคำถามบอกระดับความยากเชิงปัญญาอยู่ทางด้านซ้ายมือของข้อสอบแต่ละข้อเพื่อให้ผู้สอบใช้เป็นประโยชน์ในการทำข้อสอบ ส่วนข้อสอบอีกชุดหนึ่งจะไม่มีข้อความบอกระดับความยากเชิงปัญญา ผลการศึกษาพบว่า คะแนนรวมที่ได้จากแบบทดสอบของเพศชายกับเพศหญิงแตกต่างกัน และข้อสอบแต่ละฉบับจะจัดทำเป็น 2 ชุด ชุดแรกจะมีข้อคำถามบอกระดับความยากเชิงปัญญาอยู่ทางด้านซ้ายมือของข้อสอบแต่ละข้อเพื่อให้ผู้สอบใช้เป็นประโยชน์ในการทำข้อสอบ และข้อสอบอีกชุดหนึ่งจะไม่มีข้อความบอกระดับความยากเชิงปัญญา ปรากฏว่าชุดแรก ไม่มีผลทำให้คะแนนรวมของเพศชายและเพศหญิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่คะแนนที่ได้จากแบบทดสอบที่มีการเขียนข้อความบอกระดับความยากเชิงปัญญากับคะแนนที่ได้จากแบบทดสอบที่ไม่มีข้อความบอกระดับความยากเชิงปัญญาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

8.2 งานวิจัยภายในประเทศ

เตือนใจ เเศรษฐ์ลักโก (2511 : 45 - 48) ศึกษาการจัดระเบียบข้อทดสอบด้วยวิธีต่างๆ กันที่ส่งผลต่อนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนดาราตาม จำนวน 240 คน แบ่งเป็น 8 กลุ่ม ๆ ละ 30 คน ในวิชาภาษาอังกฤษ ภาษาไทย วิทยาศาสตร์ และ คณิตศาสตร์ โดยจัดเรียงข้อสอบด้วยวิธีที่แตกต่างกัน 4 วิธี ในแต่ละวิธีทำเป็น 2 ชุด ใช้เวลาในการสอบแตกต่างกัน และมีแบบทดสอบทั้งหมด 8 ชุด คือ ชุดที่ 1 เรียงแบบทดสอบจากง่ายไปยาก ชุดที่ 2 เรียงแบบทดสอบเหมือนชุดที่ 1 แต่เพิ่มเวลาสอบอีกฉบับละ 10 นาที ชุดที่ 3 เรียงแบบทดสอบจากวิชาคณิตศาสตร์ วิทยาศาสตร์ ภาษาไทย และ ภาษาอังกฤษ ในแต่ละฉบับเรียงข้อสอบจากง่ายไปยาก ชุดที่ 4 เรียงเหมือนชุดที่ 3 แต่เพิ่มเวลาในการสอบฉบับละ 10 นาที ชุดที่ 5 เรียงแบบทดสอบจากฉบับภาษาอังกฤษ ภาษาไทย วิทยาศาสตร์และคณิตศาสตร์ในแต่ละฉบับเรียงข้อสอบแบบสลับ ชุดที่ 6 เรียงเหมือน ชุดที่ 5 แต่เพิ่มเวลาในการสอบอีกฉบับละ 10 นาที ชุดที่ 7 เรียงแบบผสม จัดทำเป็น 5 ฉบับย่อยแต่ละฉบับย่อยจะมีข้อสอบทุกวิชา ชุดที่ 8 เรียงเหมือนชุดที่ 7 แต่เพิ่มเวลาในการสอบอีกฉบับละ 9 นาที โดยวิธีแมชกรุป (Matched Group) ผลการวิจัยพบว่า การเรียงข้อสอบต่างกัน มีผลทำให้ค่าความเชื่อมั่นของข้อทดสอบวิชาภาษาอังกฤษแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยค่าความเชื่อมั่นที่เรียงข้อทดสอบแบบผสมมีค่าต่ำสุด ส่วนการเรียงข้อทดสอบที่แตกต่างกัน มีผลทำให้ค่าความเชื่อมั่นของข้อทดสอบวิชาคณิตศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยความเชื่อมั่นของข้อทดสอบที่เรียงแบบผสมมีค่าต่ำสุด

นิโลบล นิมกังรัตน์ (2519 : 56 - 58) ได้ศึกษาอิทธิพลของการจัดเรียงลำดับข้อคำถาม ที่จำแนกตามสมรรถภาพสมองที่มีต่อคุณภาพของแบบทดสอบคณิตศาสตร์และภาษาไทยในระดับชั้นมัธยมศึกษาปีที่ 3 กับทุกโรงเรียนในเขตเทศบาลจังหวัดนครสวรรค์ ปีการศึกษา 2518 จำนวน 540 คน ผลการศึกษาพบว่า การเรียงข้อคำถามต่างกันทำให้ค่าความเชื่อมั่น ความยากมาตรฐาน และผลการสอบของนักเรียนในแบบทดสอบภาษาไทยแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าความเที่ยงตรงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่ค่าอำนาจจำแนกไม่แตกต่างกัน ส่วนแบบทดสอบวิชาคณิตศาสตร์ การจัดเรียงลำดับข้อคำถามต่างๆ กันมีผลทำให้คุณภาพของแบบทดสอบและผลการสอบของนักเรียนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และสำหรับแบบทดสอบภาษาไทยชุดที่มีการเรียงลำดับข้อคำถามแบบวัดความรู้ ความจำ ความเข้าใจ การนำไปใช้หรือสูงกว่าการนำไปใช้ มีความเหมาะสมกับกลุ่มนักเรียนที่มีความถนัดทางการเรียนสูงและต่ำ ส่วนแบบทดสอบคณิตศาสตร์ชุดที่มีการเรียงข้อคำถาม แบบวัดความรู้ความจำ ความเข้าใจ การนำไปใช้ หรือสูงกว่าการนำไปใช้นั้น ข้อคำถามแบบวัดความรู้ความจำมีความเหมาะสมกับกลุ่มนักเรียนที่มีความถนัดทางการเรียนสูงและต่ำ สำหรับกลุ่มนักเรียนที่มีความถนัดทางการเรียนปานกลาง จะใช้แบบทดสอบที่เรียงด้วยวิธีใดก็ได้

ศศิธร สุวรรณสุข (2519 : 72 - 90) ศึกษาการจัดเรียงลำดับข้อคำถามที่ต่างกันจะส่งผลต่อค่าความเที่ยงตรง และค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ เครื่องมือที่ใช้เป็นแบบทดสอบวิชาคณิตศาสตร์ 3 ฉบับ คือ ทักษะทางคณิตศาสตร์ การแก้ปัญหาทางคณิตศาสตร์ การหาเหตุผลทางคณิตศาสตร์ โดยมีการจัดเรียงลำดับข้อคำถามที่แตกต่างกัน 4 วิธี ได้แก่ การเรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อยากไปหาข้อง่าย เรียงตามหัวข้อเนื้อหาวิชา และเรียงจากข้อง่ายไปหาข้อยากในแต่ละหัวข้อเนื้อหาวิชา กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 ปีการศึกษา 2518 สังกัดองค์การ

บริหารส่วนจังหวัดสมุทรปราการ จำนวน 200 คน แบ่งเป็นนักเรียนที่มีความสามารถทางคณิตศาสตร์สูง และนักเรียนที่มีความสามารถทางคณิตศาสตร์ต่ำ แล้วสุ่มนักเรียนในแต่ละระดับเข้ากลุ่ม ๆ ละ 25 คน ผลจากการศึกษาพบว่า การเรียงลำดับข้อคำถามต่างกัน มีผลทำให้ค่าความเที่ยงตรงของแบบทดสอบวัดการแก้ปัญหาทางคณิตศาสตร์ และแบบทดสอบวัดการหาเหตุผลทางคณิตศาสตร์ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นในแบบทดสอบวัดการแก้ปัญหาทางคณิตศาสตร์ และค่าความเชื่อมั่นในแบบทดสอบวัดการหาเหตุผลทางคณิตศาสตร์แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติและการเรียงลำดับข้อคำถามแตกต่างกัน มีผลทำให้คะแนนจากแบบทดสอบวัดทักษะทางคณิตศาสตร์ และคะแนนจากแบบทดสอบวัดการแก้ปัญหาทางคณิตศาสตร์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่คะแนนที่ได้จากแบบทดสอบวัดการหาเหตุผลทางคณิตศาสตร์แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สิริพร พุ่มแก้ว (2521 : 30 - 39) ศึกษาผลการเรียงลำดับข้อสอบต่างกัน 4 แบบ คือ เรียงจากง่ายไปยาก เรียงจากยากมาง่าย เรียงตามเนื้อหา และ เรียงแบบสุ่ม โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน เครื่องมือที่ใช้แบบทดสอบผลสัมฤทธิ์วิชาภาษาไทย และวิชาคณิตศาสตร์ ผลการศึกษาพบว่า การจัดเรียงข้อสอบทั้ง 4 แบบ มีผลต่อค่าความเที่ยงตรงเล็กน้อย คะแนนการสอบวิชาภาษาไทยที่จัดเรียงตามเนื้อเรื่องแตกต่างจากการจัดเรียงแบบอื่น คะแนนการสอบวิชาคณิตศาสตร์ที่จัดเรียงจากง่ายไปยาก แตกต่างจากแบบยากไปหาง่าย และคะแนนการสอบที่จัดเรียงตามเนื้อเรื่อง แตกต่างจากการเรียงแบบสุ่ม

แห่งน้อย เพชรรัตน์ (2521 : 36 - 40) ศึกษาว่าการเรียงลำดับข้อคำถามของแบบทดสอบด้วยวิธีต่างกันที่มีผลต่อคุณภาพของแบบทดสอบวิชาวิทยาศาสตร์ คือ เรียงจากง่ายไปยาก เรียงตามลำดับความยากแบบหมุนเวียน และแบบสุ่ม โดยมีได้คำนึงถึงเนื้อหาของข้อคำถามนั้นว่าจะมีผลต่อคุณภาพของแบบทดสอบวิชาวิทยาศาสตร์ และแบบทดสอบทักษะกระบวนการทางวิทยาศาสตร์อย่างไร กลุ่มตัวอย่างที่ศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2520 ในเขตจังหวัดนครปฐม จำนวน 300 คน ผลจากการศึกษาพบว่า การเรียงลำดับข้อคำถามแบบสุ่มในแบบทดสอบทักษะกระบวนการทางวิทยาศาสตร์ส่งผลต่อความเชื่อมั่นสูงกว่าการเรียงคำถามแบบอื่น ๆ ส่วนการเรียงลำดับข้อคำถามที่เรียงจากง่ายไปยาก เรียงตามลำดับความยากแบบหมุนเวียน ส่งผลต่อความเที่ยงตรง ความเชื่อมั่น ความยากง่าย และอำนาจจำแนกในแบบทดสอบทั้ง 2 ฉบับไม่แตกต่างกัน

จินตนา ท่อทรัพย์สินชัย (2533 : 46 - 49) ได้ศึกษาความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ที่เรียงจากข้อง่ายที่สุดไปหาข้อยากที่สุด จำนวน 50 ข้อ โดยการตรวจให้คะแนนโดยใช้กฎซึลลิ่ง ที่กำหนดจุดซึลลิ่งต่างกัน 5 จุด และตรวจให้คะแนนโดยการไม่กำหนดจุดซึลลิ่ง ผลการศึกษาพบว่า การตรวจให้คะแนนโดยกำหนดจุดซึลลิ่งจุดที่ 5 เท่านั้น ทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้น เมื่อเทียบกับวิธีการตรวจแบบไม่กำหนดจุดซึลลิ่ง ส่วนการตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 1,2,3 และ 4 ให้ค่าความเชื่อมั่นต่ำกว่าการตรวจที่กำหนดจุดซึลลิ่งจุดที่ 5 และการตรวจแบบไม่กำหนดจุดซึลลิ่งอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเที่ยงตรง พบว่าการตรวจแบบไม่กำหนดจุดซึลลิ่งให้ค่าความเที่ยงตรงสูงสุด

สวงศ์ ร่มเงิน (2541 : 57 - 59) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ที่มีรูปแบบการจัดเรียงข้อคำถามตามค่าความยากทางสถิติ และ พฤติกรรมการเรียนรู้ต่างกัน จำนวน 8 ฉบับ ๆ ละ 40 ข้อ กับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จังหวัดเชียงใหม่ จำนวน 1,600 คน

ผลการศึกษาพบว่า

1. ความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ ที่เรียงข้อคำถามตามค่าความยากจากค่าสูงไปค่าต่ำจะมีค่าความเชื่อมั่นสูงกว่า เรียงข้อคำถามตามค่าความยากจากค่าต่ำไปค่าสูง และเมื่อทดสอบความแตกต่างพบว่าต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ ที่เรียงข้อคำถามตามพฤติกรรมจากขั้นต่ำไปขั้นสูง มีค่าความเชื่อมั่นสูงกว่าที่เรียงข้อคำถามตามพฤติกรรมจากขั้นสูงไปขั้นต่ำ และเมื่อทดสอบความแตกต่างพบว่าต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่เรียงตามพฤติกรรมจากขั้นต่ำไปขั้นสูงและเรียงข้อคำถามภายในแต่ละพฤติกรรมตามค่าความยากจากค่าต่ำไปค่าสูง มีค่าความเชื่อมั่นสูงสุด และเมื่อทดสอบความแตกต่างพบว่ามีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ศิริวิมล สุขสมบูรณ์ (2541 : 69) ศึกษาผลการวิเคราะห์แบบแผนการตอบข้อสอบของแบบทดสอบเลือกตอบที่มีรูปแบบการจัดเรียงข้อคำถามต่างกัน โดยเรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อง่ายไปหาข้อยากตามเนื้อหาบทเรียนเรียงตามพฤติกรรมที่ต้องการวัด กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 2 ปีการศึกษา 2540 สถาบันเทคโนโลยีราชมงคลในเขตกรุงเทพมหานครที่ทำการสอนในสาขาวิชาบริหารธุรกิจ จำนวน 1,276 คน เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก วิชาเศรษฐศาสตร์เบื้องต้น จำนวน 40 ข้อ เกี่ยวกับความรู้เบื้องต้นทางเศรษฐศาสตร์ ระบบเศรษฐกิจ การผลิต และการบริโภคกลุ่ม ผลการศึกษาพบว่าค่าความเชื่อมั่นของแบบทดสอบด้วยวิธีแบ่งครึ่งฉบับมีค่าตั้งแต่ 0.6270 ถึง 0.8253 โดยแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถามแบบเรียงจากข้อง่ายไปหาข้อยากมีค่าความเชื่อมั่นต่ำสุดและแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถาม แบบเรียงจากข้อง่ายไปหาข้อยากตามเนื้อหาของบทเรียนมีค่าความเชื่อมั่นสูงสุด ส่วนค่าความเชื่อมั่นของแบบทดสอบด้วยวิธีสอบซ้ำมีค่าตั้งแต่ 0.5530 ถึง 0.6856 โดยแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถามแบบเรียงจากข้อง่ายไปหาข้อยากมีค่าความเชื่อมั่นต่ำสุด และแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถามตามพฤติกรรมที่ต้องการวัดมีค่าความเชื่อมั่นสูงสุด ส่วนค่าความเชื่อมั่นของแบบทดสอบทั้ง 3 ฉบับที่คำนวณด้วยวิธีแบ่งครึ่งฉบับและวิธีสอบซ้ำมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บังอร กมลวัฒนา (2542 : 56 - 58) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบ และการจัดเรียงปัญหาต่างกัน โดยศึกษากับนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 สังกัดคณะกรรมการการศึกษาเอกชน กรุงเทพมหานครจำนวน 1,600 คน เครื่องมือที่ใช้เป็นแบบทดสอบวัดความสามารถในการแก้ปัญหา ตามกระบวนการแก้ปัญหาของเวียร์ (Weir) จำนวน 4 ฉบับ คือ 1) แบบทดสอบเลือกตอบที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก ชนิด 5 ตัวเลือก 10 สถานการณ์ จำนวน 40 ข้อ 2) แบบทดสอบเลือกตอบที่เรียงข้อปัญหาโดยวิธีสุ่มชนิด 5 ตัวเลือก 10 สถานการณ์ จำนวน 40 ข้อ 3) แบบทดสอบแบบตอบสั้นเรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก 10 สถานการณ์ จำนวน 40 ข้อ 4) แบบทดสอบแบบตอบสั้นที่เรียงข้อปัญหาโดยวิธีสุ่ม 10 สถานการณ์ จำนวน 40 ข้อ ผลการศึกษาพบว่าค่าความเชื่อมั่นแบบเลือกตอบที่เรียงข้อปัญหาตามความยากมาตรฐานเฉลี่ยจากง่ายไปยาก กับแบบเลือกตอบที่เรียงข้อปัญหาโดยวิธีสุ่มแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และค่าความเชื่อมั่นที่จัดเรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก

แบบเลือกตอบ กับการจัดเรียงข้อปัญหาตามความยากมาตรฐานเฉลี่ยจากง่ายไปยากแบบตอบสนองที่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

รัชนี้ กันภัย (2543 : 77 - 78) ศึกษาค่าความเชื่อมั่น และค่าความเที่ยงตรงของแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ที่สอบด้วยไมโครคอมพิวเตอร์ จำนวน 4 แบบ ๆ ละ 50 ข้อโดยแบบทดสอบแต่ละแบบจะมีคำถามเหมือนกันทุกข้อค่าความเชื่อมั่นของแบบทดสอบที่มีการจัดเรียงข้อสอบ และการกำหนดเวลาให้ทำข้อสอบวิเคราะห์โดยใช้สัมประสิทธิ์ r_B เท่ากับ 0.88, 0.92, 0.81 และ 0.89 ตามลำดับ

ในการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ ทดสอบความแตกต่างด้วยสูตร UX_1 พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($UX_1 = 12.39, P < .01$) และเมื่อพิจารณาเป็น รายคู่พบว่า

1. ค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบตามค่าความยากโดยกำหนดเวลาให้ทำข้อสอบเป็นรายข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบตามค่าความยากโดยกำหนดเวลาให้ทำข้อสอบรวมทั้งฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ค่าความเชื่อมั่นของแบบทดสอบ ที่จัดเรียงข้อสอบแบบสุ่มโดยกำหนดเวลาให้ทำข้อสอบเป็นรายข้อกับค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบแบบสุ่มโดยกำหนดเวลาให้ทำข้อสอบรวมทั้งฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำเป็นรายข้อเมื่อจัดเรียงข้อสอบตามค่าความยากกับค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำเป็นรายข้อ เมื่อจัดเรียงข้อสอบแบบสุ่มมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4. ค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำรวมทั้งฉบับ เมื่อจัดเรียงข้อสอบตามค่าความยาก กับค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำรวมทั้งฉบับเมื่อจัดเรียงข้อสอบแบบสุ่มมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากเอกสารและงานวิจัยดังกล่าวข้างต้นแสดงให้เห็นว่า วิธีการตรวจให้คะแนน การจัดเรียงข้อคำถามมีผลต่อค่าความเชื่อมั่นของแบบทดสอบ ผู้วิจัยจึงมีความสนใจที่จะศึกษาวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกัน 2 แบบ คือแบบการตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อหนึ่ง และการตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อหนึ่งมากกว่า 1 ตัว จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับការไม่กำหนดตำแหน่งการตอบผิดของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีการเรียงข้อคำถามตามค่าความยากจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด (ค่า $P .80 - .20$) ว่าจะมีค่าความเที่ยงตรง และค่าความเชื่อมั่นแตกต่างกันหรือไม่และการตอบผิดตำแหน่งใดมีผลต่อค่าความเชื่อมั่นของแบบทดสอบสูงสุด

9. สมมติฐานในการศึกษาค้นคว้า

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าแตกต่างกัน
2. ค่าความเชื่อมั่นระหว่างตำแหน่ง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
3. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกัน
4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สองมีค่าแตกต่างกัน
5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษา



ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยต้องการศึกษาคำความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่เรียงข้อคำถามตามคำความยากจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด ที่ตรวจให้คะแนนโดยกฎข้อที่ต่างกันสองแบบ ซึ่งมีวิธีดำเนินการศึกษาตามรายละเอียด ดังนี้

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 16 โรงเรียน

การเลือกกลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 8 โรงเรียน จำนวน 438 คน โดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขั้นตอนดังนี้

1. ประมาณกลุ่มตัวอย่างที่ใช้ในการศึกษาด้วยความเชื่อมั่น 95% ($\alpha = .05$) เมื่อเทียบจากตาราง ขนาดของกลุ่มตัวอย่างที่ควรเลือกจากประชากร พบว่าต้องใช้กลุ่มตัวอย่างอย่างน้อย 286 คน (ล้วน สายยศ และ อังคณา สายยศ . 2540 : 400 ,อ้างอิงมาจาก Yamane.1967) แต่ในการศึกษาค้นคว้านี้ ใช้กลุ่มตัวอย่าง 438 คน

2. แบ่งโรงเรียนออกเป็นชั้นโดยใช้ขนาดโรงเรียนเป็นชั้น (Strata) และโรงเรียนเป็นหน่วยการสุ่ม โดยแบ่งโรงเรียนออกเป็น 3 ขนาด คือ โรงเรียนขนาดใหญ่ โรงเรียนขนาดกลาง และโรงเรียนขนาดเล็ก สุ่มโรงเรียนมาประมาณ 50% ของแต่ละขนาด ได้โรงเรียนขนาดใหญ่ 2 โรงเรียน โรงเรียนขนาดกลาง 4 โรงเรียน และโรงเรียนขนาดเล็ก 2 โรงเรียน โดยใช้นักเรียนที่สุ่มได้เป็นกลุ่มตัวอย่างดังแสดงในตาราง

ตาราง 9 จำนวนนักเรียนที่เป็นกลุ่มตัวอย่างในการวิจัย จำแนกตามขนาดโรงเรียน

โรงเรียน	รวม
<u>ขนาดใหญ่</u>	
1. โรงเรียนวัดชัยพฤกษมาลา	104
2. โรงเรียนวัดปากน้ำฝ่งเหนือ	82
<u>ขนาดกลาง</u>	
1. โรงเรียนวัดมะกอก	40
2. โรงเรียนวัดโพธิ์ (ราษฎร์ผดุงผล)	64
3. โรงเรียนวัดราษฎร์ฐาน	64
4. โรงเรียนวัดตลิ่งชัน	40
<u>ขนาดเล็ก</u>	
1. โรงเรียนวัดไก่อี๋ย (สวัสดิ์ประชาชนกุล)	26
2. โรงเรียนวัดเกาะ	18
รวม	438

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

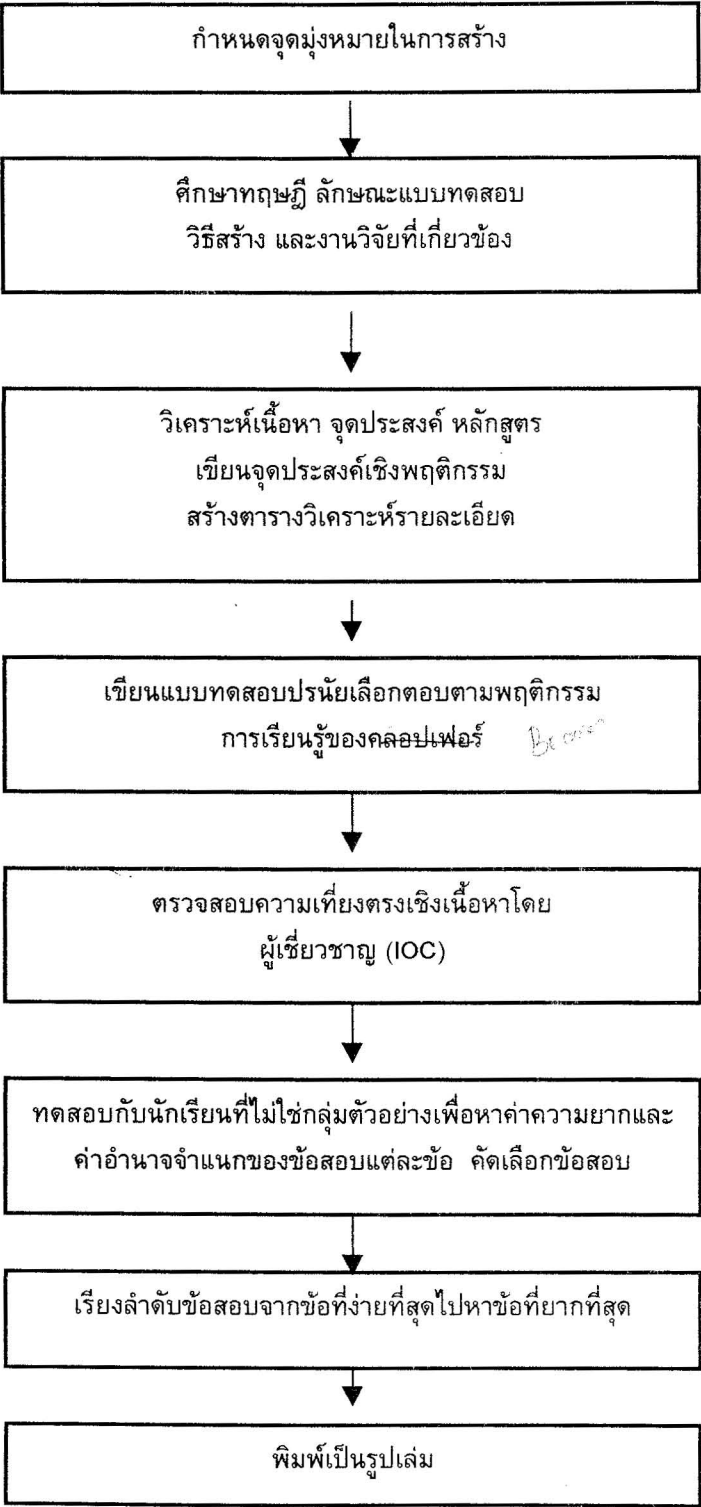
เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ชนิดเลือกตอบ 4 ตัวเลือก วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ซึ่งผู้วิจัยเป็นผู้สร้างขึ้นเอง ตามพฤติกรรมการเรียนรู้ของคลอปเฟอร์ จำนวน 50 ข้อ แบ่งเป็นหัวข้อเรื่องได้ดังนี้

- 1. สสารและความร้อน
- 2. แสง
- 3. ไฟฟ้า
- 4. แรง แรงดัน ความกดดัน
- 5. สารเคมี

วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า



เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้มีขั้นตอนการสร้าง คือ



รายละเอียดขั้นตอนการสร้างเครื่องมือ

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบปรนัยแบบเลือกตอบวิชาวิทยาศาสตร์ที่จัดเรียง

ข้อคำถามตามค่าความยากง่าย

2. ศึกษาทฤษฎี เอกสาร ตำรา ลักษณะแบบทดสอบ วิธีสร้าง และงานวิจัยที่เกี่ยวข้อง

3. ศึกษาหลักสูตรวิชาวิทยาศาสตร์ประถมศึกษาตอนปลาย พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) วิเคราะห์เนื้อหา จุดประสงค์การเรียนรู้ใน ป.02 และคู่มือครูการสอนวิชาวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) สร้างตารางวิเคราะห์ที่กำหนดจำนวนข้อสอบ

ตาราง 10 วิเคราะห์รายละเอียดวิชา วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม
	ความรู้ ความจำ	ความ เข้าใจ	กระบวนการสืบเสาะหา ความรู้วิทยาศาสตร์	การนำ ความรู้ไปใช้	จำนวน ข้อ
1. สสารและความร้อน					
1.1 จำแนกสสารและการเปลี่ยนแปลงสถานะ ของสสารได้	1	-	2	1	4
1.2 อธิบายและจำแนกการส่งผ่านความร้อน ด้วยวิธีต่างๆได้	1	2	1	1	5
1.3 อธิบายปรากฏการณ์เกี่ยวกับการ เปลี่ยนแปลงของสสารในปรากฏการณ์ ธรรมชาติและสามารถนำความรู้ไปใช้ ประโยชน์ในการดำรงชีวิตได้	1	-	1	1	3
2. แสง					
2.1 อธิบายการเดินทางของแสงผ่าน ตัวกลางต่างๆได้	1	2	-	-	3
2.2 อธิบายและทดลองเกี่ยวกับการเดินทาง ของแสง การเกิดเงา และนำหลักการใน เรื่องนั้นๆมาใช้ประโยชน์	-	3	-	2	5
2.3 บอกความหมาย อธิบายกฎการสะท้อน ของแสงและกฎการหักเหของแสงได้	1	-	-	-	1
2.4 อธิบายส่วนประกอบของดวงตาที่ เกี่ยวกับการมองเห็นได้	1	1	-	-	2

ตาราง 10 (ต่อ)

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม จำนวน ข้อ
	ความรู้ ความจำ	ความ เข้าใจ	กระบวนการสืบเสาะหา ความรู้วิทยาศาสตร์	การนำ ความรู้ไปใช้	
3. ไฟฟ้า					
3.1 อธิบายและทดลองเกี่ยวกับแหล่งกำเนิดไฟฟ้าได้	1	1	-	2	4
3.2 อธิบายและจำแนกวงจรไฟฟ้า ตัวนำไฟฟ้า และฉนวนไฟฟ้าได้	-	1	-	-	1
3.3 อธิบายวิธีใช้ วิธีป้องกันอันตรายจาก อุปกรณ์ไฟฟ้า และใช้ไฟฟ้าอย่างประหยัดได้	1	1	-	1	3
4. แรง แรงดัน ความกดดัน					
4.1 อธิบายปรากฏการณ์ที่เกี่ยวข้องกับ ความกดดันของอากาศ ความกดดันน้ำ และความชื้นของอากาศได้	1	1	1	-	3
4.2 อธิบายและทดลองเกี่ยวกับการใช้เครื่อง มือที่นำความรู้ในเรื่องแรงดัน ความดันอากาศ และน้ำได้ถูกต้อง	1	1	1	-	3
4.3 อธิบายการเปลี่ยนแปลงของแรงดัน ความดันของอากาศและน้ำ ความชื้นของ อากาศ ผลกระทบที่มีต่อการดำรงชีวิตของ มนุษย์	1	1	-	1	3
5. สารเคมี					
5.1 อธิบายและทดลองเกี่ยวกับการ เปลี่ยนแปลงของสารเคมีอย่างง่ายได้	2	-	2	1	5
5.2 บอกส่วนประกอบ วิธีการทดสอบและ อันตรายของสารเคมีที่ใช้ในชีวิตประจำวันได้	-	-	1	2	3
5.3 บอกวิธีใช้และวิธีเก็บรักษายาฆ่าแมลง สารกำจัดศัตรูพืชได้อย่างปลอดภัย	-	2	-	-	2
รวมจำนวนข้อ	13	16	9	12	50

4. สร้างแบบทดสอบปรนัยแบบเลือกตอบชนิด 4 ตัวเลือก โดยให้สอดคล้องกับตารางกำหนดรายละเอียดในข้อ 3 จำนวน 100 ข้อ

5. ให้ผู้เชี่ยวชาญจำนวน 5 ท่านตรวจสอบแบบทดสอบเป็นรายข้อว่า วัดได้ตรงตามพฤติกรรมการเรียนรู้หรือไม่ โดยใช้เกณฑ์การพิจารณาความเห็นของผู้เชี่ยวชาญ (ค่า IOC) ตั้งแต่ .60 ขึ้นไปถือว่าเป็นข้อสอบที่วัดได้ตรงตามพฤติกรรม (บุญเชิด ภิญญอนันตพงษ์.2545 : 95) นำไปใช้ทดสอบได้ ซึ่งผลจากการตรวจสอบ ปรากฏว่าแบบทดสอบมีค่า IOC ตั้งแต่ .60 – 1.00 อยู่ในเกณฑ์ที่ใช้ได้ทั้งหมดแต่ต้องนำมาปรับปรุงภาษาที่ใช้ให้รัดกุมและเหมาะสมในบางข้อ จากนั้นผู้วิจัยนำแบบทดสอบจำนวน 100 ข้อไปใช้ในการทดสอบค่าความยากและค่าอำนาจจำแนกกับนักเรียนที่สุ่มมาสำหรับการสร้างเครื่องมือ

ตัวอย่าง แบบประเมินความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

คำชี้แจง ให้ท่านพิจารณาว่า ข้อคำถามแต่ละข้อต่อไปนี้วัดได้ตรงตามพฤติกรรมการเรียนรู้ของคลอปเฟอร์จริงหรือไม่ โดยให้พิจารณานำหนัก แล้วใส่เครื่องหมาย ✓ ลงในช่องคะแนน

- +1 เมื่อท่านพิจารณาแล้วเห็นว่า ข้อคำถามข้อนี้วัดได้ตรงตามพฤติกรรมข้อนี้จริง
- 0 เมื่อท่านพิจารณาแล้วไม่แน่ใจว่า ข้อคำถามนี้วัดได้ตรงตามพฤติกรรมข้อนี้จริง
- 1 เมื่อท่านพิจารณาแล้วเห็นว่า ข้อคำถามข้อนี้วัดได้ไม่ตรงตามพฤติกรรมข้อนี้จริง

ตัวอย่าง การประเมินความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

พฤติกรรมการเรียนรู้ที่ ต้องการวัด	ข้อสอบ	คะแนน		
		+1	0	-1
ความรู้ความจำ	(0) เมื่อใส่ก้อนหินลงในแก้วน้ำปรากฏว่าระดับน้ำสูงขึ้นแสดงถึงคุณสมบัติข้อใดของสสาร ก. สสารมีตัวตน ข. สสารมีน้ำหนัก ค. สสารต้องการที่อยู่ ง. สสารสามารถสัมผัสได้	✓		
ความเข้าใจ	(00) สถานการณ์ใดต่อไปนี้เป็นหลักการเดียวกับการปล่อยยางรัดลูกโป่งแล้วลูกโป่งเคลื่อนที่ ก. การเตะลูกบอลให้เคลื่อนที่ ข. การวางลูกบอลบนพื้นเอียงแล้วปล่อยให้กลิ้งลงมา ค. การพายเรือด้วยไม้พายทำให้เรือเคลื่อนที่ไปข้างหน้า ง. การปล่อยฟองอากาศในน้ำและฟองอากาศลอยขึ้นสู่ผิวน้ำ	✓		

พฤติกรรมการเรียนรู้ที่ ต้องการวัด	ข้อสอบ	คะแนน		
		+1	0	-1
กระบวนการสืบเสาะหา ความรู้ทางวิทยาศาสตร์	(000) ข้อมูลใดที่ใช้เครื่องมือชนิดเดียวกันวัดได้โดยไม่ต้อง คำนวณ ก. ความลึกของบ่อน้ำ และปริมาตรของถังน้ำ (ข.) ขนาดของรอบเอว และความยาวของห้องเรียน ค. เส้นผ่านศูนย์กลางของลูกฟุตบอล และเส้นผม ง. ความยาวของโต๊ะเรียน และความหนาของกระดาษ	✓		
การนำไปใช้	(0000) ถ้านำปั๊มน้ำตั้งไฟจนเดือดยกลงจากเตาไฟ ปิดฝาแล้วเอาน้ำราด นักเรียนจะสังเกตเห็นอะไร (ก.) ปั๊มน้ำจะบวมขึ้น ข. ปั๊มน้ำจะพองออก ค. ปั๊มน้ำจะแตกออกเป็นสองส่วน ง. ฝาปั๊มน้ำจะกระเด็นหลุดออกมา	✓		

6. นำแบบทดสอบที่คัดเลือกไว้ ไปทดสอบกับนักเรียนกลุ่มที่สุ่มมาสำหรับการสร้างเครื่องมือ
จำนวน 200 คน ตรวจให้คะแนนแบบ 0 - 1 แล้วนำไปวิเคราะห์เพื่อหาคุณภาพของข้อสอบรายข้อในด้าน
ความยากง่าย และค่าอำนาจจำแนก โดยใช้โปรแกรม CTIA ซึ่งพัฒนาโดย สุวัฒน์ สุขมลสันต์ ผลการ
วิเคราะห์พบว่า แบบทดสอบมีค่าความยากตั้งแต่ 0.090 - 0.960 และมีค่าอำนาจจำแนกตั้งแต่
-0.241 - 0.593 หลังจากนั้นคัดเลือกข้อสอบที่มีค่าความยาก 0.20 - 0.80 และค่าอำนาจจำแนก 0.20 ขึ้น
ไปให้ครอบคลุมพฤติกรรมการเรียนรู้ และครบจำนวนข้อตามตารางวิเคราะห์ข้อสอบ โดยเป็นข้อสอบที่วัด
พฤติกรรมด้านความรู้ความจำ จำนวน 13 ข้อ ด้านความเข้าใจจำนวน 16 ข้อ ด้านกระบวนการสืบเสาะหา
ความรู้ทางวิทยาศาสตร์จำนวน 9 ข้อ และด้านการนำความรู้ไปใช้จำนวน 12 ข้อ รวมทั้งหมดจำนวน
50 ข้อ มีค่าความยาก (p) 0.250 - 0.800 และค่าอำนาจจำแนก (r) 0.222 - 0.593 (รายละเอียดปรากฏใน
ตาราง 19,20 ภาคผนวก ค)

7. จัดเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

8. จัดพิมพ์เป็นแบบทดสอบเพื่อใช้ในการวิจัย

วิธีการเก็บรวบรวมข้อมูล

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์ในการเก็บข้อมูล
2. ติดต่อกับโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง ขออนุญาตผู้บริหารโรงเรียน เพื่อทำการทดสอบระหว่างวันที่ 11 – 15 กุมภาพันธ์ 2545
3. เตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่สอบในแต่ละครั้ง
4. นำแบบทดสอบวัดผลสัมฤทธิ์ที่สร้างขึ้นไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง
5. อธิบายให้นักเรียนที่เป็นกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ และประโยชน์ที่ได้รับจากการทำแบบทดสอบ
6. ตรวจกระดาษคำตอบตามวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 2 แบบ และนำผลมาวิเคราะห์ทางสถิติ

วิธีการตรวจให้คะแนนโดยกฎชิลลิ่ง 2 แบบ

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง
2. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง

ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 แล้วกำหนดการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด ซึ่งมีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

1.1 การตอบผิดตำแหน่งที่หนึ่งให้ 1 คะแนนสำหรับข้อที่ตอบถูกต้องตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิด (0) เป็นครั้งแรก ข้อที่ตอบถูกต้องหลังจากนั้นจะไม่นำมานับคะแนน

1.2 การตอบผิดตำแหน่งที่สองให้ 1 คะแนนสำหรับข้อที่ตอบถูกต้องตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดติดต่อกันสองข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกต้องหลังจากนั้นจะไม่นำมานับคะแนน

1.3 การตอบผิดตำแหน่งที่สามให้ 1 คะแนนสำหรับข้อที่ตอบถูกต้องตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดติดต่อกันสามข้อ (000) เป็นครั้งแรก ข้อที่ตอบถูกต้องหลังจากนั้นจะไม่นำมานับคะแนน

ตัวอย่าง การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง

คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1
1.	1	14.	1	27.	1	40.	0
2.	1	15.	1	28.	1	41.	0
3.	1	16.	1	29.	1	42.	0***
4.	0*	17.	0	30.	0	43.	1
5.	1	18.	1	31.	0	44.	1
6.	0	19.	1	32.	1	45.	0
7.	1	20.	0	33.	1	46.	0
8.	0	21.	1	34.	0	47.	1
9.	1	22.	0	35.	1	48.	0
10.	1	23.	1	36.	0	49.	1
11.	1	24.	0	37.	1	50.	1
12.	0	25.	0	38.	1		
13.	0**	26.	1	39.	0		

หมายเหตุ

- * การตอบผิดตำแหน่งที่หนึ่ง นักเรียนจะได้คะแนน 3 คะแนน
 - ** การตอบผิดตำแหน่งที่สอง นักเรียนจะได้คะแนน 8 คะแนน
 - *** การตอบผิดตำแหน่งที่สาม นักเรียนจะได้คะแนน 24 คะแนน
- การไม่กำหนดตำแหน่งการตอบผิด นักเรียนจะได้คะแนน 29 คะแนน

ผลจากการกำหนดคะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องทำให้ได้คะแนน 4 ชุด คือ

- ชุดที่ 1 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่หนึ่ง
- ชุดที่ 2 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่สอง
- ชุดที่ 3 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่สาม
- ชุดที่ 4 คะแนนจากการไม่กำหนดตำแหน่งการตอบผิด

2. การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 แล้วกำหนดการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งที่การตอบผิด ซึ่งมีการตรวจให้คะแนนในแต่ละตำแหน่ง ดังนี้

2.1 การตอบผิดตำแหน่งที่หนึ่ง ให้ 1 คะแนน ตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.2 การตอบผิดตำแหน่งที่สอง ให้ 1 คะแนน ตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สอง ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.3 การตอบผิดตำแหน่งที่สาม ให้ 1 คะแนน ตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สาม ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

การไม่กำหนดตำแหน่งการตอบผิด ตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

ตัวอย่าง การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1
1.	1	14.	0	27.	1	40.	1
2.	1	15.	1	28.	1	41.	1
3.	1	16.	1	29.	1	42.	0
4.	1	17.	0	30.	1	43.	1
5.	1	18.	1	31.	0	44.	1
6.	0	19.	1	32.	1	45.	1
7.	0*	20.	0	33.	1	46.	0
8.	1	21.	1	34.	0	47.	0***
9.	0	22.	0	35.	1	48.	0
10.	1	23.	1	36.	0	49.	1
11.	1	24.	0	37.	1	50.	1
12.	1	25.	0**	38.	1		
13.	1	26.	1	39.	0		

หมายเหตุ

* การตอบผิดตำแหน่งที่หนึ่ง นักเรียนจะได้คะแนน 5 คะแนน

** การตอบผิดตำแหน่งที่สอง นักเรียนจะได้คะแนน 16 คะแนน

*** การตอบผิดตำแหน่งที่สาม นักเรียนจะได้คะแนน 31 คะแนน

การไม่กำหนดตำแหน่งการตอบผิด นักเรียนจะได้ 33 คะแนน

ผลจากการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว ทำให้ได้คะแนน

4 ชุด คือ

- ชุดที่ 1 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่หนึ่ง
- ชุดที่ 2 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่สอง
- ชุดที่ 3 คะแนนที่ได้จากข้อแรกถึงการตอบผิดตำแหน่งที่สาม
- ชุดที่ 4 คะแนนจากการไม่กำหนดตำแหน่งการตอบผิด

ขั้นตอนในการวิเคราะห์ข้อมูล

ขั้นตอนในการวิเคราะห์ข้อมูลมีการดำเนินการตามลำดับดังต่อไปนี้

1. การวิเคราะห์ข้อมูลในการทดลองเครื่องมือ (Try Out) ผู้วิจัยได้ดำเนินการดังนี้

1.1 ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

1.2 วิเคราะห์ค่าความยากโดยใช้โปรแกรม CTIA ซึ่งพัฒนาโดย สุวัฒน์ สุขมลสันต์

เพื่อจัดเรียงข้อสอบตามค่าความยากโดยเรียงจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด

2. การวิเคราะห์ข้อมูลในการวิจัย ผู้วิจัยได้ดำเนินการดังนี้

2.1 ตรวจให้คะแนนแบบทดสอบ โดยกฎชิลลิ่งที่แตกต่างกัน ทั้ง 2 แบบและนำเฉพาะข้อมูลที่

ตรวจได้ครบตามข้อตกลงเบื้องต้น มาใช้ในการวิเคราะห์ข้อมูล

2.2 หาค่าสถิติพื้นฐาน คือ

2.2.1 คะแนนเฉลี่ย

2.2.2 ส่วนเบี่ยงเบนมาตรฐาน

2.2.3 สัมประสิทธิ์การกระจาย

2.3 ประเมินค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตรสัมประสิทธิ์ r_B

(บุญเชิด ภิญญอนันตพงษ์, 2545 : 134)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐาน ได้แก่ค่าเฉลี่ย ($\bar{X}$) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) สัมประสิทธิ์การกระจาย (CV)

ของคะแนนของแบบทดสอบ

2. สถิติที่ใช้ในการหาคุณภาพเครื่องมือ

2.1 ค่าความยาก คำนวณโดยใช้สูตร (สุวัฒน์ สุขมลสันต์, 2538)

$$p = \frac{N_a}{N_t}$$

เมื่อ	N_a	แทน	จำนวนผู้สอบทั้งหมดที่ตอบตัวเลือก a
	N_t	แทน	จำนวนผู้สอบทั้งหมด
	P	แทน	ค่าความยากของตัวเลือกแต่ละข้อสำหรับผู้สอบทั้งหมด

2.2 ค่าอำนาจจำแนก คำนวณโดยสูตร (สุพัฒน์ สุกมลสันต์. 2538)

$$r_{pb} = \frac{\bar{X}_r - \bar{X}_w}{S.D._t} \sqrt{p(1-p)}$$

เมื่อ	r_{pb}	แทน	ดัชนีค่าอำนาจจำแนก
	$\bar{X}_r$	แทน	ค่าเฉลี่ยของคะแนนผู้สอบที่ตอบข้อสอบนั้นถูก
	$\bar{X}_w$	แทน	ค่าเฉลี่ยของคะแนนผู้สอบที่ตอบข้อสอบนั้นผิด
	$S.D._t$	แทน	ค่า S.D. ของคะแนนรวมของผู้สอบทั้งหมด
	P	แทน	อัตราส่วนของผู้ที่ตอบข้อสอบข้อนั้นถูก

2.3 ค่าความเที่ยงตรงเชิงพินิจ (Face Validity) ของแบบทดสอบคำนวณโดยสูตร

(บุญเชิด ภิญโญนนต์พงษ์. 2545 : 95)

$$IOC = \Sigma R / N$$

เมื่อ	IOC	แทน	ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์
	R	แทน	ความคิดเห็นของผู้เชี่ยวชาญ
	ΣR	แทน	ผลรวมของคะแนนจากผู้เชี่ยวชาญในแต่ละข้อ
	N	แทน	จำนวนผู้เชี่ยวชาญทั้งหมด

3. ประเมินค่าความเชื่อมั่นของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกันทั้ง

2 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่ง โดยใช้สูตรสัมประสิทธิ์ r_B (บุญเชิด ภิญโญนนต์พงษ์. 2545 : 134)

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

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

$$\lambda_1 = [\Sigma B_{1i} - P_1(\bar{X} - 1)] / S_x^2$$

$$\lambda_2 = [\Sigma B_{2i} - P_2(\bar{X} - 1)] / S_x^2$$

⋮

$$\lambda_K = [\Sigma B_{Ki} - P_K(\bar{X} - 1)] / S_x^2$$

$$\bar{X} = \sum P_i$$

B_{ji} แทน สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใดๆ ถูกทั้งคู่

K แทน จำนวนข้อสอบ

4. ทดสอบนัยสำคัญของความแตกต่างของความเชื่อมั่นในแต่ละแบบ ใช้สูตร UX_1
(บุญเชิด ภิญโญนนตพงษ์. 2543 : 57 : อ้างอิงจาก Woodruff & Felde. 1986: 393 – 413)

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

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

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

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

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

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

$$\bar{n}_h = \frac{M}{\sum \frac{1}{n_i}}$$

UX_1 แทน สถิติทดสอบที่มีการแจกแจงแบบ χ^2 df = m - 1

r แทน ค่าความเชื่อมั่นที่คำนวณได้

r_{ij} แทน ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร

S_u^2 แทน ค่าความแปรปรวนของ u_i

c_u แทน ค่าความแปรปรวนของ u_j

M แทน จำนวนค่าสัมประสิทธิ์ความเชื่อมั่น

N แทน จำนวนนักเรียนที่ทดสอบ

$\bar{u}$ แทน ค่าเฉลี่ยของคะแนน

$\bar{n}_h$ แทน ค่าเฉลี่ยค่าสังเกต

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้เข้าใจตรงกันในการแปลความหมายผลของการวิเคราะห์ข้อมูล ผู้วิจัยจึงกำหนดสัญลักษณ์และอักษรย่อสำหรับการวิเคราะห์ข้อมูล ดังนี้

$\bar{X}$	แทน	คะแนนเฉลี่ยทั้งฉบับ
S.D.	แทน	ส่วนเบี่ยงเบนมาตรฐานทั้งฉบับ
CV	แทน	สัมประสิทธิ์การกระจาย
แบบที่ 1	แทน	กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง
แบบที่ 2	แทน	กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว
r_B	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
UX_1	แทน	สถิติทดสอบที่มีการแจกแจงแบบ χ^2
df	แทน	ชั้นของความเป็นอิสระ (Degrees of Freedom)
**	แทน	มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลนำเสนอเป็น 4 ตอน คือ

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

ตอนที่ 2 การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่ง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

3.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

3.2 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกัน ทั้งสองแบบจากการตอบผิดตำแหน่งเดียวกันของแบบทดสอบ จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม

รายละเอียดของผลการวิจัย มีดังนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกัน ทั้งสองแบบ

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนผลการสอบของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา วิทยาศาสตร์ที่เรียงตามค่าความยากจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด จำนวน 50 ข้อ ที่ตรวจให้คะแนน โดยกฎชิลลิ่งทั้งสองแบบ คือ การตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนด ตำแหน่งการตอบผิดของกลุ่มตัวอย่าง มาวิเคราะห์หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์ การกระจาย (CV) ปรากฏผลดังแสดงในตาราง 11

ตาราง 11 ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้งสองแบบจากการ ตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

วิธีการตรวจให้ คะแนน ตำแหน่ง การตอบผิด	กฎชิลลิ่งแบบคะแนน ศูนย์ต่อเนื่อง			กฎชิลลิ่งแบบคะแนนศูนย์ ต่อ เนื่อง มากกว่า 1 ตัว		
	$\bar{X}$	S.D.	CV	$\bar{X}$	S.D.	CV
ตำแหน่งที่หนึ่ง	3.98	2.99	75.13	9.05	5.48	60.55
ตำแหน่งที่สอง	9.05	5.48	60.55	11.75	6.51	55.40
ตำแหน่งที่สาม	15.50	7.27	46.90	15.40	7.20	47.87
การไม่กำหนดตำแหน่ง	23.01	5.51	23.95	23.01	5.51	23.95

ผลการวิเคราะห์ตามตาราง11 ปรากฏว่า

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 3.98 – 23.01 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่ง มีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.95 – 75.13 โดยสัมประสิทธิ์การกระจายของ การตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่ หนึ่งมีการกระจายสูงสุด

2. การตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 9.05 – 23.01 โดยการตรวจให้คะแนนแบบการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.95 – 60.55 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

ตอนที่ 2 การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้งสองแบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนผลการสอบของแบบทดสอบมาตรวจให้คะแนนโดยกฎซึลลิ่งทั้งสองแบบจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมาหาค่าความยากแต่ละข้อ ค่าความยากทั้งฉบับ สัดส่วนของผู้ตอบข้อสอบสองข้อใด ๆ ถูกทั้งสองข้อ ค่าแลมด้า (λ_i) ค่าความแปรปรวนของคะแนนทั้งฉบับ แล้วนำค่าที่ได้ ไปแทนค่าในสูตรการหาค่าสัมประสิทธิ์ r_B (บุญเชิด ภิญญ โอนันตพงษ์, 2538 : 51) เพื่อหาค่าความเชื่อมั่นของแบบทดสอบ ปรากฏผลดังแสดงในตาราง 12

ตาราง 12 ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้งสองแบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

ค่าความเชื่อมั่น	กฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง			กฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว			การไม่กำหนดตำแหน่งการตอบผิด
	ตำแหน่งที่ 1	ตำแหน่งที่ 2	ตำแหน่งที่ 3	ตำแหน่งที่ 1	ตำแหน่งที่ 2	ตำแหน่งที่ 3	
r_B	0.907	0.897	0.882	0.897	0.901	0.888	0.678

ผลการวิเคราะห์ตามตาราง 12 ปรากฏว่า

1. การตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สามมีค่าความเชื่อมั่นตั้งแต่ 0.882 - 0.907 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุด

2. การตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นตั้งแต่ 0.888 - 0.901 โดยการตอบผิดตำแหน่งที่สอง มีค่าความเชื่อมั่นสูงสุด

3. การตรวจให้คะแนนโดยไม่กำหนดตำแหน่งการตอบผิดมีค่าความเชื่อมั่น เท่ากับ 0.678

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นของคะแนนแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

ในการเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดผู้วิจัยได้นำค่าความเชื่อมั่นมาเปรียบเทียบกันและทดสอบความแตกต่าง ดังนี้

3.1 เปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1)

ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิดมาเปรียบเทียบกัน โดยใช้สูตร UX_1 ปรากฏผลดังตาราง 13

ตาราง 13 การเปรียบเทียบค่าความเชื่อมั่นของคะแนนแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

สูตร	ตำแหน่งที่ หนึ่ง	ตำแหน่งที่ สอง	ตำแหน่งที่ สาม	การไม่ กำหนด ตำแหน่ง	UX_1
r_B	0.907	0.897	0.882	0.678	260.271 **

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 13 ปรากฏว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด ที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดอย่างน้อย 1 คู่ มีค่าแตกต่างกัน ดังนั้น จึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ ด้วยสูตร UX_1 อีกครั้ง ปรากฏผลดังตาราง 14

ตาราง 14 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบ
ศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

r_B	ตำแหน่งการตอบผิด	ตำแหน่ง ที่ หนึ่ง	ตำแหน่ง ที่ สอง	ตำแหน่ง ที่ สาม	การไม่ กำหนด ตำแหน่ง
0.907	ตำแหน่งที่หนึ่ง	-	1.875	7.973 **	185.762 **
0.897	ตำแหน่งที่สอง		-	3.653	198.131 **
0.882	ตำแหน่งที่สาม			-	220.722 **
0.678	การไม่กำหนดตำแหน่ง				-

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 14 ปรากฏว่าค่าความเชื่อมั่นรายคู่ของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เกือบทุกตำแหน่งการตอบผิด ยกเว้นการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง และการตอบผิดตำแหน่งที่สองกับการตอบผิดตำแหน่งที่สาม แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.2 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นของคะแนนแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 (แบบที่ 2)

ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมาเปรียบเทียบกัน โดยใช้สูตร UX_1 ปรากฏผลดังตาราง 15

ตาราง 15 การเปรียบเทียบค่าความเชื่อมั่นของคะแนนแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

สูตร	ตำแหน่งที่ หนึ่ง	ตำแหน่งที่ สอง	ตำแหน่งที่ สาม	การไม่ กำหนด ตำแหน่ง	UX_1
r_B	0.897	0.901	0.888	0.678	397.752 **

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 15 ปรากฏว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด ที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด อย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้น จึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ ด้วยสูตร UX_1 อีกครั้ง ปรากฏผลดังตาราง 16

ตาราง 16 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบ ศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนด ตำแหน่งการตอบผิด

r_B	ตำแหน่งการตอบผิด	ตำแหน่ง ที่ หนึ่ง	ตำแหน่ง ที่ สอง	ตำแหน่ง ที่ สาม	การไม่ กำหนด ตำแหน่ง
0.897	ตำแหน่งที่หนึ่ง	-	0.469	3.921	198.131 **
0.901	ตำแหน่งที่สอง		-	8.156 **	268.500 **
0.888	ตำแหน่งที่สาม			-	279.286 **
0.678	การไม่กำหนดตำแหน่ง				-

** มีนัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 16 ปรากฏว่าค่าความเชื่อมั่นรายคู่ของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เกือบทุก ตำแหน่งการตอบผิด ยกเว้นการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง ตำแหน่งที่หนึ่งกับตำแหน่งที่สาม มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน ทั้งสองแบบของแบบทดสอบจากการกำหนดการตอบผิดตำแหน่งเดียวกัน

การวิเคราะห์ในตอนนี้ ผู้วิจัยนำค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนที่แตกต่างกัน ทั้งสองแบบ จากการตอบผิดตำแหน่งเดียวกันมาเปรียบเทียบ และทดสอบความแตกต่าง โดยใช้สูตร UX_1 ปรากฏผลดังตาราง 17

ตาราง 17 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน ทั้ง 2 แบบของแบบทดสอบ จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม

	ค่าความเชื่อมั่นจากวิธีการตรวจให้คะแนน		UX_1
	แบบที่ 1	แบบที่ 2	
r_B ตำแหน่งที่หนึ่ง	0.907	0.897	1.875
r_B ตำแหน่งที่สอง	0.897	0.901	0.468
r_B ตำแหน่งที่สาม	0.882	0.888	0.884

ผลการวิเคราะห์ตามตาราง 17 ปรากฏว่า ค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สามจากการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 2 แบบ มีค่าแตกต่างกัน อย่างไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

การศึกษาวิจัยครั้งนี้ เป็นการศึกษาค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน 2 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด และค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 2 แบบจากการตอบผิดตำแหน่งเดียวกันว่าจะมีค่าความเชื่อมั่นแตกต่างกันหรือไม่ ซึ่งสรุปสาระสำคัญได้ดังนี้

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาค้นคว้าในครั้งนี้มีจุดมุ่งหมายหลัก

1. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง

2. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

3. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

4. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง มากกว่า 1 ตัวจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 2 แบบจากการตอบผิดตำแหน่งที่หนึ่ง

6. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 2 แบบจากการตอบผิดตำแหน่งที่สอง

7. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 2 แบบจากการตอบผิดตำแหน่งที่สาม

สมมติฐานในการศึกษาค้นคว้า

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
2. ค่าความเชื่อมั่นระหว่างตำแหน่ง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
3. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกัน
4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สองมีค่าแตกต่างกัน
5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกัน

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 16 โรงเรียน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2544 ของโรงเรียนสังกัดกรุงเทพมหานคร สำนักงานเขตตลิ่งชัน จำนวน 8 โรงเรียน จำนวน 438 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 6 แบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก 1 ฉบับ จำนวน 50 ข้อ ซึ่งผู้วิจัยสร้างขึ้นตามพฤติกรรมการเรียนรู้ของคลอปเฟอร์ และมีการเรียงข้อคำถามตามค่าความยากโดยเรียงข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

สรุปผลการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้มุ่งที่จะศึกษาค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิง แต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกัน และค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 2 แบบ จากการตอบผิดตำแหน่งเดียวกันมีค่าความเชื่อมั่นแตกต่างกัน ผลการวิจัยมีดังนี้

1. ค่าสถิติพื้นฐาน

1.1 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 3.98 - 23.01 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.95 - 75.13 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

1.2 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 9.05 - 23.01 โดยการตรวจให้คะแนนแบบการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุด และการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.95 - 60.55 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

2. ค่าความเชื่อมั่น

2.1 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สามมีค่าความเชื่อมั่นตั้งแต่ 0.882 - 0.907 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุด

2.2 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นตั้งแต่ 0.888 - 0.901 โดยการตอบผิดตำแหน่งที่สอง มีค่าความเชื่อมั่นสูงสุด

2.3 การตรวจให้คะแนนโดยไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นต่ำสุด เท่ากับ 0.678

3. การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงทั้ง 2 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

3.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ.01 ได้แก่ การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม การตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่ง การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่ง การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่ง ส่วนการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สอง การตอบผิดตำแหน่งที่สองกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่เป็นนัยสำคัญทางสถิติ

3.2 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ได้แก่ การตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการตอบผิดตำแหน่งที่สาม การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด ส่วนการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สอง การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกัน ทั้ง 2 แบบของแบบทดสอบจากการตอบผิดตำแหน่งเดียวกัน

4.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกัน
ทั้ง 2 แบบของการตอบผิดตำแหน่งที่หนึ่ง มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.2 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกัน
ทั้ง 2 แบบของการตอบผิดตำแหน่งที่สอง มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกัน
ทั้ง 2 แบบของการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

จากการวิเคราะห์ข้อมูล พบว่า

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดตำแหน่งที่หนึ่ง กับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการศึกษาของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon . 1987 : 52) วิลสัน และคณะ (Wilson ,et al . 1988 : 216) จินตนา ต่อทรัพย์สินชัย (2533 : 47) มานะ คิ้วเจริญ (2539 : 60) และพิมพ์ร ภูมิรา (2544 : 103) เมื่อพิจารณาถึงความแตกต่างพบว่า มีส่วนที่ต่างกัน 2 ประการ ประการที่หนึ่ง คือ การตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุด สอดคล้องกับผลการศึกษาของพิมพ์ร ภูมิรา (2544 : 103) ที่พบว่าการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุด แต่แตกต่างกับผลการศึกษาของสมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon . 1987 : 52) วิลสัน และคณะ (Wilson ,et al . 1988 : 216) ที่พบว่าการตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด เหตุที่เป็นเช่นนี้อาจเป็นผลจากลักษณะของแบบทดสอบที่ใช้ต่างกัน กล่าวคือ แบบทดสอบที่ผู้วิจัยใช้สร้างขึ้นตามพฤติกรรมการเรียนรู้ของคลอปเฟอร์ ประกอบด้วยความรู้ด้านความรู้ ความจำ ความเข้าใจ กระบวนการเสาะหาความรู้ทางวิทยาศาสตร์ การนำความรู้และวิธีการวิทยาศาสตร์ไปใช้ ส่วนแบบทดสอบของสมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon . 1987) วิลสัน และคณะ (Wilson ,et al . 1988) ใช้แบบทดสอบมาตรฐานที่แบ่งเป็นฉบับย่อย ฉบับละพฤติกรรม นอกจากนั้นสูตรการคำนวณค่าความเชื่อมั่นที่ต่างกัน ทำให้เกิดความแตกต่างกัน กล่าวคือ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon . 1987) วิลสัน และคณะ (Wilson ,et al . 1988) หาค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์แอลฟา (Alpha

Coefficient) ส่วนการศึกษาวิจัยครั้งนี้ผู้วิจัยหาค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์ r_B (Coefficient r_B) (บุญเชิด ภิญญอนันตพงษ์. 2538 : 5)

2. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 2) จากการตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด และการตอบผิดตำแหน่งที่สองกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการศึกษาของสมบุญ สิริวงค์ (Suriyawongse, Somboon .1987 : 52) วิลสันและคณะ (Wilson , et al.1988 : 216) จินตนา ต่อทรัพย์สินชัย (2533 : 47) มานะ คิ้วเจริญ (2539 : 60) และพิมพ์ ภูมิรา (2544 : 108) เมื่อพิจารณาพบว่าการตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุดสอดคล้องกับผลการวิจัยของสมบุญ สิริวงค์ (Suriyawongse, Somboon . 1987 : 52) วิลสัน และคณะ (Wilson ,et al .1988) ที่พบว่า การตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุดส่วนการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สอง การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้อาจเป็นเพราะว่าในการตรวจให้คะแนนตามกฎชิลลิ่งพบว่าเมื่อนักเรียนเริ่มทำผิดที่จุดชิลลิ่งที่หนึ่งแล้วคะแนนที่ได้หลังจากนั้นเกิดจากการเดามากกว่าการใช้ความรู้จริงในการตอบ จึงทำให้ค่าความเชื่อมั่นในจุดดังกล่าวแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง ได้แก่ การตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่อง และการตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้เนื่องมาจากการค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งทั้งสองแบบ จากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกันน้อยมาก เนื่องจากเงื่อนไขการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว ตำแหน่งที่หนึ่ง ก็คือตำแหน่งเดียวกันกับการตอบผิดตำแหน่งที่สองของการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องจึงมีค่าความเชื่อมั่นเท่ากันเท่ากับ 0.879 และจากการทดสอบเป็นรายคู่ของการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แสดงให้เห็นว่าค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งของวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้งสองแบบมีค่าไม่แตกต่างกัน

4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้งสองแบบ ของแบบทดสอบจากการตอบผิดตำแหน่งที่สอง ได้แก่ การตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่อง และการตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้อาจเนื่องมาจากค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวมีค่าไม่แตกต่างกัน

5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้งสองแบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สาม ได้แก่ การตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่อง และการตรวจให้คะแนนแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้อาจเนื่องมาจากยิ่งทำข้อสอบมากข้อสอบก็ยิ่งยาก เพิ่มมากขึ้น และโอกาสในการตอบผิดก็เพิ่มมากขึ้นตามเงื่อนไขการตรวจให้คะแนนที่เปิดโอกาสให้ตอบผิด ติดต่อกันได้มากขึ้นทั้ง 2 แบบ จึงทำให้ค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สามของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 2 แบบ มีค่าใกล้เคียงกัน

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้และในการวิจัยดังต่อไปนี้

1. ข้อเสนอแนะในการนำไปใช้

1.1 จากการวิจัยครั้งนี้พบว่า ในการสอบคัดเลือกที่มีการประเมินผลผู้เข้าสอบจำนวนมาก และต้องการใช้ผลการสอบในการจัดตำแหน่งมากกว่าการพิจารณาความรอบรู้ในเนื้อหาของผู้สอบ ควรใช้วิธีการตรวจให้คะแนนโดยกฎซัลลิงแบบคะแนนศูนย์ต่อเนื่อง ตำแหน่งที่หนึ่ง เพราะให้ค่าความเชื่อมั่นสูงสุด และช่วยประหยัดเวลาในการตรวจให้คะแนน

1.2 การนำวิธีการตรวจให้คะแนนโดยกฎซัลลิงไปใช้ตรวจให้คะแนนในการจัดการเรียนการสอน ตามปกติในชั้นเรียน ต้องพิจารณาถึงความเหมาะสม เนื่องจากวิธีนี้ช่วยประหยัดเวลาในการตรวจให้คะแนน สามารถช่วยจัดการได้ แต่ไม่สามารถวัดความรู้ความสามารถของนักเรียนได้ครอบคลุมเนื้อหา

2. ข้อเสนอแนะในการวิจัย

2.1 เนื่องจากวิธีการตรวจให้คะแนนโดยกฎซัลลิงเป็นวิธีการหนึ่งที่ช่วยจัดการได้ จึงควรทำการวิจัยในทำนองนี้ โดยเพิ่มตัวแปรของกลุ่มตัวอย่างที่มีความสามารถต่างกัน จำแนกเป็นกลุ่มเก่ง กลุ่มปานกลาง กลุ่มอ่อน หรือในระดับชั้นอื่นๆ

2.2 ควรศึกษาค่าความเชื่อมั่นและความเที่ยงตรงเมื่อใช้แบบทดสอบที่มีจำนวนข้อต่างกัน และใช้วิธีการตรวจให้คะแนนโดยกฎซัลลิง

2.3 ควรศึกษาวิธีการตรวจให้คะแนนโดยกฎซัลลิงกับกลุ่มวิชาอื่น เพื่อนำผลมาเปรียบเทียบกับการศึกษาครั้งนี้

บรรณานุกรม

บรรณานุกรม

- เกริกชัย ฮาบเจริญ. (2525). การเปรียบเทียบการตอบ และการตรวจให้คะแนนแบบใหม่ และวิธีทดสอบความมั่นใจ และแบบบรรณานุกรมที่มีต่อค่าความเชื่อมั่น และคะแนนการเดาของแบบทดสอบเลือกตอบ.
- ปริญญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- โกวิท ประวาลพุกษ์ และสมศักดิ์ สินธุระเวชญ์. (2523). การประเมินผลในชั้นเรียน. กรุงเทพฯ ฯ : วัฒนาพานิช.
- จิตรกร ตั้งเกษมสุข. (2525). ชุดตำราเรียนความเป็นมาของการศึกษาไทย. กรุงเทพฯ ฯ : บางกอกการพิมพ์.
- จินตนา ต่อทรัพย์สินชัย. (2533). ผลของการตรวจให้คะแนน โดยใช้กฎซิลลิงที่มีต่อค่าความเชื่อมั่น และค่าความเที่ยงตรง ของแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์. ปริญญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ชวาล แพรัตกุล. (2516). เทคนิคการวัดผล. พิมพ์ครั้งที่ 5. วัฒนาพานิช.
- _____. (2517). การทดสอบเพื่อค้น และพัฒนาสมรรถภาพ. หน้า 46 - 47. กรุงเทพฯ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2518). เทคนิคการวัดผล. พิมพ์ครั้งที่ 6. กรุงเทพฯ ฯ : วัฒนาพานิช.
- ชลินทร์ ศิลปสมศักดิ์. (2544). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีการตรวจให้คะแนนตามกฎซิลลิง โดยมีการจัดเรียงข้อสอบและกำหนดจุดซิลลิงต่างกัน. ปริญญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ต่าย เชียงฉิน. (2526) ทฤษฎีการทดสอบและการวัดผลการศึกษา. ภาควิชาประเมินผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่.
- เตือนใจ เศรษฐลักโก. (2511). การศึกษาถึงการจัดระเบียบข้อสอบด้วยวิธีต่าง ๆ ที่จะส่งผลต่อความสามารถของนักเรียนชั้น ป.7. ปริญญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ทวี ทองคำ. (2526). การเปรียบเทียบค่าความเที่ยง ความตรง และอำนาจจำแนกของแบบทดสอบชนิดเลือกตอบที่ใช้คำสั่งและวิธีการให้คะแนนที่ต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ: บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ธงชัย ชิวปรีชา และคณะ. (2526). เอกสารการสอนชุดวิชาการสอนวิทยาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช : หน่วยที่ 11 ปฏิบัติการวิทยาศาสตร์และหน่วยที่ 12 การวัดและประเมินผลการเรียนการสอนวิทยาศาสตร์. กรุงเทพฯ ฯ : โรงพิมพ์ยูไนเต็มนัดปรัดกชั้น.

- ธีรศักดิ์ อินทรมาตย์. (2520). อิทธิพลของวิธีการตอบแบบทดสอบปรนัยชนิดเลือกตอบวิธีต่าง ๆ ที่มีต่อค่าความเชื่อมั่น ความเที่ยงตรงและปริมาณการเดา. ปรียญานิพนธ์ กศ.ม. การวัดผลการศึกษา).
- กรุงเทพ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- นภา กาญจนกิจโสภณ. (2528). การเปรียบเทียบคะแนนสอบระหว่างการให้น้ำหนักคะแนนรายข้อเท่ากันตาม ทฤษฎีมาตรฐานเดิมกับการให้น้ำหนักคะแนนรายข้อต่างกันตาม ทฤษฎีการตอบสนองข้อคำถาม. วิทยานิพนธ์ ค.ม. (ภาควิชาชีพการศึกษา). กรุงเทพ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- นิโลบล นิ่มกิ่งรัตน์. (2519). อิทธิพลของการจัดเรียงลำดับข้อคำถามที่จำแนกตามสมรรถภาพสมองที่มีผลต่อคุณภาพของแบบทดสอบคณิตศาสตร์ และภาษาไทย ม. 3. ปรียญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- แน่นน้อย เพชรรัตน์. (2521). ผลการเรียงลำดับข้อคำถามด้วยวิธีต่างกันที่มีผลต่อคุณภาพของแบบทดสอบวิชา ทักษะกระบวนการวิทยาศาสตร์. ปรียญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- บังอร กมลวัฒนา. (2542). การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน. ปรียญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- บุญชม ศรีสะอาด. (2540). การวิจัยทางการวัดผลและประเมินผล. กรุงเทพ ฯ : สุวีริยาสาส์น
- บุญธรรม กิจปรีดาบริสุทธิ์. (2535). คู่มืออาจารย์ : การวัดและการประเมินผลการเรียนการสอน = Handbook : measurement and evaluation in teaching. กรุงเทพ ฯ : บี แอนด์ บี พับลิเคชั่น.
- บุญเชิด ภิญญอนันตพงษ์. (2521). การวัดและประเมินผลการศึกษา : ทฤษฎีและการประยุกต์. คณะศึกษาศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2526). การสอนแบบอิงเกณฑ์ : แนวคิดและวิธีการ. กรุงเทพ ฯ : ภาควิชาพื้นฐานการศึกษา. คณะศึกษาศาสตร์. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2533. พฤษภาคม - สิงหาคม). "Congeneric Part Reliability," การวัดผลการศึกษา. 12 (34) : 28 - 31.
- _____. (2536. มกราคม - เมษายน). "Congeneric Model," การวัดผลการศึกษา. 11(42) : 8 - 20 .
- _____. (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์. ปรียญานิพนธ์ กศ.ด. (การวัดผลการศึกษา). กรุงเทพ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- _____. (2538). "แบบทดสอบคะแนนจริงสัมพันธ์สัมพันธ์ r_B ," วารสารการวัดผลการศึกษา. กรุงเทพ ฯ : ม.ป.ท.
- _____. (2542). สัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน. กรุงเทพ ฯ : ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2545). ประมวลสาระชุดวิชาการพัฒนาเครื่องมือสำหรับการประเมินการศึกษา(หน่วยที่ 3). กรุงเทพ ฯ : มหาวิทยาลัยสุโขทัยธรรมาธิราช. สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช.

- เบญจวรรณ แยมละมูล.(2538).การศึกษาความเที่ยงตรงแบบหลายลักษณะหลายวิธีของแบบทดสอบวัดเจตคติที่มีต่อโรงเรียนและครูในระดับประถมศึกษา. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ๕ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- พจน์ สะเพียรชัย. (2515). “ ทฤษฎีการตรวจให้คะแนนกับการเดา ,” พัฒนาวัดผล 8. หน้า 21 - 25. กรุงเทพฯ ๕ : อักษรสัมพันธ์.
- พินชา สวัสดิ์มงคล. (2530). การเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบในการวัดระดับความรู้ที่ต่างกันวิธีโดยให้คะแนนต่างกันในกลุ่มผู้สอบที่มีผลสัมฤทธิ์ทางการเรียนต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พินิจ อุไรรักษ์. (2533). ผลของวิธีการให้คะแนนที่มีต่อคะแนนสอบ ค่าความเชื่อมั่น และค่าความเที่ยงตรง. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พิมพร ภูมิรา.(2544).การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกัน 3 แบบ.ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา).กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- เพ็ญศรี สว่างเนตร. (2520). ความเที่ยงตรงของแบบสอบสองชนิดเลือกตอบด้วยเทคนิคการให้คะแนนที่ต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ๕ :บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- เพลินพิศ คุณคำ. (2532). การประมาณความเชื่อมั่นและปริมาณการเดาของแบบทดสอบปรนัยเลือกตอบที่มีวิธีการตอบและการตรวจให้คะแนนวิธี 0 - 1, วิธีของคูมบีส และ วิธีของอนันต์. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ๕ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ภพ เลหาไพบุลย์.(2542).แนวการสอนวิทยาศาสตร์.พิมพ์ครั้งที่ 3.กรุงเทพฯ : ไทยวัฒนาพานิช. มหาวิทยาลัย, ทบวง. (2524). ชุดการเรียนการสอนสอนสำหรับครูคณิตศาสตร์. 335 หน้า.
- มานะ ถั่วเจริญ. (2539). การศึกษาผลการตรวจให้คะแนนโดยกฎชิลลิ่งที่กำหนดจุดชิลลิ่งต่างกันของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยาก และ พฤติกรรมการเรียนรู้. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ๕ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- มิ่ง เทพครเมือง.(2539).การตรวจสอบความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบการคิดเอกละเอียดด้านสัญลักษณ์หกฉบับ ตามทฤษฎีของกิลฟอร์ด. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ๕ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- เยาวลักษณ์ น้อยนาแสง. (2535). การเปรียบเทียบคุณภาพของข้อสอบความถนัดทางการเรียนแบบปรนัยชนิดเลือกตอบซึ่งเป็นผลจากวิธีการตอบและการตรวจให้คะแนนที่แตกต่างกัน. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ๕ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.

- รัชนี้ กันภัย. (2543). ผลของการจัดเรียงข้อสอบต่างกันและกำหนดเวลาให้ทำต่างกันที่มีต่อค่าความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบที่สอบด้วยไมโครคอมพิวเตอร์. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ล้วน สายยศ . (2517). “ การวัดผลสัมฤทธิ์วิชาคณิตศาสตร์,” เอกสารการประชุมวิชาการครั้งที่ 2. หน้า 36 - 39. กรุงเทพฯ : โรงพิมพ์ศาสนา.
- ล้วน สายยศ . (2526). “ วิเคราะห์ผลการเดาข้อสอบก่อนสอบและหลังสอบกับความสามารถจริง,” วารสารการวัดผลการศึกษา. กรุงเทพฯ .
- ล้วน สายยศและอังคณา สายยศ. (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ : ชมรมเด็ก.
- _____. (2540). สถิติวิทยาทางการวิจัย. พิมพ์ครั้งที่ 3. กรุงเทพฯ : สุวีริยาสาส์น.
- วรรณวดี ม้าลำพอง.(2532).การสร้างแบบทดสอบความถนัด.วิทยาลัยครูเชียงใหม่.ม.ป.ท.
- วัชรภรณ์ จิตราต. (2538). การเปรียบเทียบคุณภาพแบบทดสอบเลือกตอบชนิดถูกผิดหลายตัวเลือกที่มีจำนวนตัวเลือกและวิธีตรวจให้คะแนนต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศศิธร สุวรรณสุข. (2519). การศึกษาค่าความเที่ยงตรง ค่าความเชื่อมั่นของแบบทดสอบที่เรียงลำดับคำถามต่าง ๆ กันในวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 7. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศิริวิมล สุขสมบูรณ์. (2541). ผลการวิเคราะห์แบบแผนการตอบข้อสอบของแบบทดสอบเลือกตอบที่มีรูปแบบการจัดเรียงข้อคำถามต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศึกษาธิการ, กระทรวง. (2533 ก). คู่มือหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : โรงพิมพ์การศาสนา.
- _____. (2533 ข). หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : โรงพิมพ์การศาสนา.
- _____. (2534). คู่มือการประเมินผลการเรียนตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2534). คู่มือครูคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6. หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กระทรวงศึกษาธิการ : โรงพิมพ์คุรุสภา.
- สงวต ร่มเงิน. (2541). การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ที่มีรูปแบบการจัดเรียงข้อคำถามตามค่าความยากทางสถิติและพฤติกรรมการเรียนรู้ต่างกัน. ปรินญาณพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- สิริพร พุ่มแก้ว. (2521). ผลของการเรียงลำดับข้อสอบต่อค่าความเที่ยงตรงของแบบทดสอบและผลสัมฤทธิ์ทางการเรียน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

- สุพัฒน์ สุกมลสันต์. (2528.พฤษภาคม - สิงหาคม).“ ผลการกำหนดน้ำหนักคะแนนต่างคุณภาพของแบบสอบ ” ,
วารสารวิธีวิทยาการวิจัย. กรุงเทพฯ ฯ : จุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2528).การเปรียบเทียบผลของวิธีการให้คะแนนต่อค่าความเที่ยง ความตรง และความคงที่
ของแบบทดสอบการอ่าน ความเข้าใจภาษาอังกฤษที่มีการสร้างความรู้ที่ต่างกัน. วิทยานิพนธ์. จ.ม.
(การวัดและประเมินผลการศึกษา) . กรุงเทพฯ ฯ :จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร
- สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ. กอง. วิชาการ.(2534). รายงานการประเมินความก้าวหน้า
คุณภาพนักเรียนชั้นประถมศึกษาปีที่ 6 ระดับประเทศ ปีการศึกษา 2532. กรุงเทพมหานคร :
อักษรไทย.
- _____. (2538). การวัดและประเมินผลในชั้นเรียนกลุ่มทักษะคณิตศาสตร์. เอกสารกองวิชาการ ลำดับที่ 18 /
2538. กรุงเทพฯ ฯ : โรงพิมพ์คุรุสภาลาดพร้าว.
- สำนักงานคณะกรรมการการศึกษาแห่งชาติ.สำนักงานรัฐมนตรี.(2540). แผนพัฒนาการศึกษาแห่งชาติฉบับที่ 8.
(พ.ศ. 2540 - 2544). กรุงเทพฯ ฯ : อรรถผลการพิมพ์.
- สำราญ มีแจ้ง. (2525). ผลของคำสั่งการใช้คะแนนที่ต่างกันต่อค่าความเที่ยง ความตรง และอำนาจจำแนกของ
แบบทดสอบชนิดเลือกตอบ. วิทยานิพนธ์ จ.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ :
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- สำเริง บุญเจริญรัตน์. (2525). “การพัฒนาทฤษฎีลาเท้นเพื่อวิเคราะห์ข้อสอบ,” วารสารการวัดผลการศึกษา 4 (2) :
1 – 13 ; กันยายน – สิงหาคม.
- _____. (2529). ทฤษฎีการวัดและการประเมินผลการศึกษา. พิมพ์ครั้งที่ 2 กรุงเทพฯ ฯ : สยามศึกษา.
- สมพงษ์ อิศวฤกษ์. (2542). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจและความยาวของ
แบบทดสอบต่างกัน. ปรินญาณพนธ์ กศ.ม.(การวัดผลการศึกษา). กรุงเทพฯ ฯ :
บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- อนันต์ ศรีโสภ. (2520) . การวัดผลและประเมินผล. กรุงเทพฯ ฯ : ไทยวัฒนาพานิช.
- _____. (2524). การวัดผลและการประเมินผลการศึกษา. พิมพ์ครั้งที่ 2. กรุงเทพฯ ฯ : ไทยวัฒนาพานิช.
- _____. (2526). “เราจะพัฒนาวิธีการตอบและตรวจให้คะแนนข้อสอบแบบเลือกตอบได้อย่างไร,” สามัญศึกษา.
13 - 19 : เมษายน.
- _____. (2532). การทดลองรูปแบบใหม่ของข้อสอบแบบเลือกตอบ. กรุงเทพฯ ฯ : คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- อัมพร ยิ้มยิ้ม. (2533). ผลของการตรวจให้คะแนน โดยกฎซิลลิงที่กำหนดจุดซิลลิงต่างกันที่มีต่อค่าการเดา และ
ลักษณะโค้งอินเฟอร์เมชันของแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์. ปรินญาณพนธ์ กศ.ม.
(การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
ถ่ายเอกสาร.
- อัมพร วิชัยศรี.(2541).การแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้างและการประมาณค่าความเชื่อมั่นรวมของ
แบบทดสอบความสามารถทางสมองตามแนวโครงสร้างของโอติส – เลนอน. ปรินญาณพนธ์ กศ.ม.
(การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
ถ่ายเอกสาร.
- อรนุช ศรีสะอาด.(2526).คู่มือการสอนวิชาการวัดผลการศึกษา.ภาควิชาพื้นฐานของการศึกษา มหาวิทยาลัย
ศรีนครินทรวิโรฒ มหาสารคาม.

- อรรณณ ตันท์เจริญรัตน์ . (2517). *การศึกษาวิธีการตอบ และการตรวจให้คะแนนแก่ข้อสอบ ปรนัยแบบเลือกตอบ*. ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย ศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- อวยพร ธิบุญญากาญจน์. (2526). *การเปรียบเทียบผลการวิเคราะห์แบบสอบอุปมาอุปมัยด้วยคลาสสิกอลโมเดลกับราชคโมเดล*. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- อุทุมพร จามรมาน. (2536). *ข้อสอบ : การสร้างและการพัฒนา*. กรุงเทพฯ ฯ : ฟีนีฟับลิชชิง.
- Abu - Sayf, F.X. (1975, December) "Reliability Effectiveness of the Conventional Formula Score," *Journal of Educational Resarch*. 69 (4) ; 160 - 162 .
- Abu - Sayf, F.X. and James J. Diamond. (1976, Novenber-December). " Effect of Confidence Level in Multiple - Choice Test Answer on Reliability and Validity of Score," *Journal of Educational Resarch*. 70 : 62 .
- Adams, Georgia S. *Measurement and Evaluation in Educational Psychology and Guidance*. New york : Holt Rinehart and Winston Inc., 1964. 654 p. Mehrens, W.A & Lehman, I.J. _____.(1964). *Measurement and Evaluation in Educational Psychology*. 3rd ed. Japan : CBS College Publishing.
- Albert, Wilfred George. (1970 , October). "Scoring for Partial Knowledge in Mathematics Testing : A study of Modication of and Extention of Multiple - Choice Items Applied to the Testing of Achievement in Mathematics ," *Dersertation Abstracts International*. 31 (4) : 1619 A .
- Allen, M.J. and Yen, W.M (1979). *Introduction to measurement theory*. Monterey : Brooks / Cole Publishing Co.
- Anastasi, Anne. (1968). *Psychological Testing*. 3 rd ed. London : Macmilan Company.
- Bloom, B.S. and others (Eds.) (1956). *Taxonomy of Educational Objective : Handbook 1. Cognitive Domain*. New York : Mckay.
- Campbell, D.T. and Fiske .(1959, April). "Convergent and Discriminant Validity by the Multitrait - Multimethod ," *Psychological Bullentin* p.56.
- Coombs ,Clydes H, Milholand , J.E. and Womer, F.B. (1956). " The Assessment of Partial Knowledge," *Journal of Educational and Psychological Measurement* . 16 (1) : 13 - 17 ; Spring.
- Davis, Federick B. (1966). *Educational Measurement and Their Interpretation*. Wadsworth Publishing Co.
- Davis Frederick B. and Gordon, Fiffer. (1959). " The Effect on test Reliability and validity of Scoring Aptitude and Achievement test with Weight for Every Choice," *Journal of Educational and Psychological Measurement* . 19 : 159 - 170 ; Summer.
- Dressel, P.L and Schmid, J. (1953). " Some Modification of the Multiple - Choice Item," *Journal of Educational and Psychological Measurement* . 13 : 474 - 495.
- Ebel, Robert L . (1965). *Measuring Educational Achievement*. New Jursy : Prentice - Hall.

- Ebel, Robert L. (1965) "Confidence Weighting and Test Reliability," *Journal Of Education Measurement*. 2 : 49 – 57.
- Edwards, A.L. (1957). *Techniques of attitude Scale construction*. New York : Appletion –Century - Crofts.
- Ferguson, G. A. & Takane Yoshio. (1989). *Statistical Analysis in Psychology and Education*. (6 th ed.) McGraw - Hill Book Company.
- Flangher, R. Melton, R. and Myers, C. (1968). " Item Rearrangement under Typical Test Conditional," *Educational and Psychological Measuerment*. p 813 –824 .London : Sage Publications, Inc.
- Garret, Henry E. and Wood Worth, R.S. (1967). *Statistical in Psychology and Education*. David Mckay Company, Inc.
- Gibbon, Jean D., Olkin, Inqram and Sobel, Melton. (1979. September). " A Subtest Selection technique for scoring Item on a Multiple Choice test," *Psychometrika* . 44 : 259 - 270.
- Gronlund, Norman E (1976). *Measurement and Evaluation in Teaching*. (3 rd ed). New York : Macmillan.
- _____. (1977). *Constructing Achievement Tests*. 2 rd ed. Engwood Cliffs, N.J. : Prentice - Hall.
- Gulliksen, H. (1950). *Theory of mental tests*. New York : wiley.
- Hambleton, R.K. (1978). " Criterion - Reference Testing and Measurement : A Review of Technical Issues and Developments," *Reviews of Educational Research*. 48 (1) : 1 - 49, Winter - Fall.
- Jaradat, Derar and Seri Sawaged. (1986). " The Subset Selection Technique for Multiple - Choice Tests : An Empirical Inquiry," *Journal of Educational Measurement*. 23 : 396 - 375 ; Winter.
- James Diamond J. and Abu - Sayf. (1976). " Effect of Confidence Level in Multiple - Choice Test Answer on Reliability and Validity of Score," *Journal of Educational Resarch*. 70 : 62 ; November-December.
- Klosner, Naomi Certner & Gellman, Estelle Klitnick. (1973). " The Effect of Item Arrangment on classroom Test Performance : Implication for Content Validity," *Educational and Psychological Measurement*. 33 : 413 – 418. London : Sage Publications, Inc.
- Lane, S., David , J.R & Bull, Kay S. (1987). " The Effects of Knowledge of Item Arrangment, Gender, and Statistical and Cognitive Item Difficulty an Test Performance," *Educational and Psychological Measurement*. p.865 - 879. London : Sage Publications, Inc.
- Lyman, Howard B. (1963) *Test Scores and What They Mean*. Prentice-Hall,Inc., Englowood Cliff, N.J. 223. P.
- Lindquisit, E. F. (1955). *Educational Measurement*. Washington : American Council on Education.
- Lindvall, C.M. & Nitko. Anthony J. (1967). *Measuring Pupil Achievement and Aptitude*. New York : Horcourt Brace Jovanovich, Inc.
- Lord, Frederic M & Novick, Melvin R. (1968). *Statistical Theories of Mental Scores*. California : Harcourt Brace Jovanovich, Inc.
- Mehrens, William A. and Irvin Lehmann, J. (1984). *Measurement and Evaluation in Educational and Psycholigy*. New York : Rinehart and Winston.

- Mollenkopf, W. G. (1950). "An Experimental Study of the Effect on Item Analysis Data of Changing Item Placement and Test Time Limit," *Psychometrika*. 291 - 292.
- Nunnally, Jum C. (1959). *Test and Measurement*. New York : McGraw - Hill Book Company, Inc.
- _____. (1964). *Educational Measurement and Evaluation*. New York : McGraw - Hill Book Company.
- Plake, Barbara S. (1980). "Item Arrangement and Knowledge of Arrangement on Test Scores," *Journal of Experiment Education*. Washington, D.C. : JE & E. 49 : 56 - 58 ; Fall.
- Pugh , Richard C. and Brunza, J. (1975). "Effect of Confidence Weighted Scoring System on Measures of Test Reliability and Validity," *Educational and Psychological Measurement*. 35 : 73 - 78 ; Spring.
- Remmers,H.H. and Adkins, R.H. (1955). "Reliability of Multiple-Choice and True-False," *Journal Of Educational Psychology*. 33 : 385-391.
- Stanley, Julian C. and Kenncth Hopkins, D. (1972). *Educational and Psychology Evaluation*. (5 th. Ed). Englewoods Cliff ; Pretic - Hall.
- Suriyawongse, Somboon. (1987). "Acomparision of the effect of different sizes of Ceiling Rule on the Estimates of Reliability of a Mathematic Achievement Test," *Dissertation Abstracts International*. University of North Texas, 48 : 493.
- Thorndike,Robert L. *Measurement and Evaluation in Psychology and Education*. New york : John Willey and Sons, Inc., 1969.
- Torgerson, W.S. (1958). *Theory and method of Scaling*. New York : John Wiley and Son.
- Travers, Robert H.W. (1950). *How to make Achievement Test*. New York : Odyssey Press.
- Tyler, L.E. (1971). *Tests and Measurement*. (2 nd ed). Englewood Cliffs, N.J. Prentice - Hall.
- Wilson, James W. (1971). "Evaluation of Learning in Secondary School Mathematics," *Handbook on Formative and Summative Evaluation of Student Learning*. New York : McGraw - Hill Book Company.
- Wilson, Michael M. and others. (1988). "The Effects of Ceiling Rules on The International Consistency Reliability of A Mathematics Test," *Educational and Psychological Measurement* . University of North Texas, 48 : 213 – 217
- Winer, B.J. (1971). *Statistical Principle in Experimental Design*. (2 nd ed.). Tokyo : McGraw - Hill Kogaknsha.

ภาคผนวก

ภาคผนวก ก

รายนามผู้เชี่ยวชาญประเมินความสอดคล้องระหว่างข้อคำถาม
กับพฤติกรรมการเรียนรู้

รายนามผู้เชี่ยวชาญ

ผศ.ระวีวรรณ	พันธุ์พานิช	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์ชวลิต	รวยอาจิน	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์อุดมศักดิ์	นาดี	นักวิชาการศึกษา 8 กรมวิชาการ สำนักการศึกษา สังกัดกรุงเทพมหานคร
อาจารย์จันทลักษณ์	ธรรมทวีธิกุล	อาจารย์ใหญ่ โรงเรียนคริสต์ธรรมศึกษา กรุงเทพมหานคร
อาจารย์วรรณ	พิทักษ์ถิร	หมวดวิชาวิทยาศาสตร์ โรงเรียนราชินี เขตพระนคร กรุงเทพมหานคร

ภาคผนวก ข

ค่าดัชนีความสอดคล้องระหว่างข้อความกับพฤติกรรมการเรียนรู้

(Index of Congruency : IOC)

ตาราง 18 ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

ข้อ	IOC	ข้อ	IOC	ข้อ	IOC	ข้อ	IOC
1	1	26	1	51	1	76	1
2	0.8	27	1	52	1	77	1
3	1	28	1	53	1	78	1
4	0.8	29	0.8	54	1	79	0.8
5	1	30	1	55	1	80	0.8
6	1	31	1	56	1	81	0.8
7	1	32	1	57	1	82	1
8	1	33	0.8	58	1	83	1
9	1	34	1	59	1	84	1
10	1	35	1	60	1	85	1
11	1	36	1	61	0.8	86	1
12	1	37	1	62	0.6	87	1
13	1	38	1	63	1	88	1
14	1	39	1	64	1	89	1
15	1	40	1	65	1	90	0.8
16	0.8	41	1	66	1	91	0.8
17	1	42	1	67	1	92	1
18	1	43	1	68	1	93	1
19	0.6	44	1	69	0.6	94	1
20	1	45	1	70	1	95	0.6
21	1	46	1	71	1	96	0.8
22	0.6	47	1	72	1	97	1
23	0.8	48	1	73	1	98	1
24	0.8	49	1	74	0.8	99	1
25	1	50	1	75	0.8	100	1

ภาคผนวก ค

ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6

ตาราง 19 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 จากการทดลองใช้ (Try out)

ข้อที่	p	r	ข้อที่	p	r
1	0.925	0.148	26	0.250	0.315
2	0.475	0.222	27	0.630	0.167
3	0.620	0.444	28	0.680	0.481
4	0.750	0.259	29	0.770	0.352
5	0.665	0.333	30	0.390	0.056
6	0.800	0.222	31	0.570	0.333
7	0.540	0.222	32	0.555	0.352
8	0.825	0.241	33	0.865	0.167
9	0.880	0.241	34	0.590	0.370
10	0.360	0.389	35	0.435	0.074
11	0.455	0.407	36	0.770	0.333
12	0.475	0.111	37	0.665	0.315
13	0.730	0.407	38	0.550	0.370
14	0.820	0.333	39	0.650	0.352
15	0.320	0.333	40	0.555	0.056
16	0.76	0.315	41	0.570	0.389
17	0.515	0.593	42	0.810	0.315
18	0.335	0.037	43	0.500	0.352
19	0.830	0.315	44	0.495	0.296
20	0.695	0.444	45	0.460	0.093
21	0.310	0.241	46	0.645	0.407
22	0.720	0.426	47	0.495	0.296
23	0.775	0.389	48	0.695	0.463
24	0.515	-0.074	49	0.295	0.093
25	0.400	0.370	50	0.235	-0.019

ตาราง 19 (ต่อ)

ข้อที่	p	r	ข้อที่	p	r
51	0.895	0.111	76	0.835	0.241
52	0.920	0.167	77	0.555	0.296
53	0.600	0.333	78	0.495	0.241
54	0.145	-0.074	79	0.505	0.463
55	0.675	0.500	80	0.830	0.204
56	0.730	0.370	81	0.605	0.204
57	0.955	0.056	82	0.275	0.241
58	0.260	0.093	83	0.290	0.278
59	0.405	0.093	84	0.760	0.407
60	0.900	0.185	85	0.475	0.093
61	0.665	0.222	86	0.945	0.148
62	0.670	0.333	87	0.730	0.426
63	0.190	0.204	88	0.855	0.130
64	0.175	-0.074	89	0.540	0.111
65	0.835	0.167	90	0.730	0.037
66	0.315	0.315	91	0.500	0.241
67	0.900	0.130	92	0.635	0.463
68	0.415	0.352	93	0.625	0.056
69	0.730	0.278	94	0.755	0.407
70	0.150	0.000	95	0.430	-0.222
71	0.460	0.352	96	0.945	0.000
72	0.645	0.407	97	0.765	0.296
73	0.120	-0.241	98	0.510	0.241
74	0.545	0.259	99	0.185	0.167
75	0.960	0.056	100	0.090	0.037

ตาราง 20 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
 วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ใช้ในการวิจัย

ข้อที่	p	r	ข้อเดิม	ข้อที่	p	r	ข้อเดิม
1	0.800	0.222	6	26	0.555	0.296	77
2	0.770	0.352	29	27	0.555	0.352	32
3	0.765	0.296	97	28	0.550	0.370	38
4	0.760	0.407	84	29	0.545	0.259	74
5	0.755	0.407	94	30	0.540	0.222	7
6	0.730	0.426	87	31	0.515	0.593	17
7	0.730	0.407	13	32	0.510	0.241	98
8	0.730	0.370	56	33	0.505	0.463	79
9	0.720	0.426	22	34	0.500	0.352	43
10	0.695	0.463	48	35	0.500	0.241	91
11	0.695	0.444	20	36	0.495	0.296	47
12	0.680	0.481	28	37	0.495	0.241	78
13	0.670	0.333	62	38	0.475	0.222	2
14	0.665	0.333	5	39	0.460	0.352	71
15	0.665	0.222	61	40	0.455	0.407	11
16	0.650	0.352	39	41	0.415	0.352	68
17	0.645	0.407	46	42	0.400	0.370	25
18	0.645	0.407	72	43	0.390	0.560	30
19	0.635	0.463	92	44	0.360	0.389	10
20	0.620	0.444	3	45	0.320	0.333	15
21	0.605	0.204	81	46	0.315	0.315	66
22	0.600	0.333	53	47	0.310	0.241	21
23	0.590	0.370	34	48	0.290	0.278	83
24	0.570	0.389	41	49	0.275	0.241	82
25	0.570	0.333	31	50	0.250	0.315	26

ภาคผนวก ง
แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์
ชั้นประถมศึกษาปีที่ 6

**แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์
ชั้นประถมศึกษาปีที่ 6**

คำชี้แจง

1. แบบทดสอบมีทั้งหมด 1 ฉบับ จำนวน 50 ข้อ ใช้เวลา 50 นาที
2. คำถามทุกข้อเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุดเพียงข้อเดียว จากข้อ ก ข ค และ ง โดยกาเครื่องหมาย X ลงในช่องสี่เหลี่ยมใต้ข้อที่นักเรียนเลือกในกระดาษคำตอบ ดังตัวอย่างต่อไปนี้

(0) วงจรไฟฟ้าชนิดใดทำให้เครื่องใช้ไฟฟ้าทำงานได้

- ก. วงจรเปิด
- ข. วงจรปิด
- ค. วงจรตรง
- ง. วงจรสลับ

จะเห็นว่าคำตอบที่ถูกต้อง คือ ข้อ ข ให้นักเรียนกาเครื่องหมาย X ลงในช่องสี่เหลี่ยมให้ตรงกับข้อ

(0) ได้ตัวเลือก ข ในกระดาษคำตอบดังนี้

ข้อ	ก	ข	ค	ง
0		X		

3. ถ้านักเรียนต้องการเปลี่ยนคำตอบใหม่ให้ขีด (=) ทับรอบเดิมก่อนแล้ว เลือกตอบใหม่ เช่น ต้องการเปลี่ยนคำตอบจากข้อ ข เป็นข้อ ง ให้ทำดังนี้

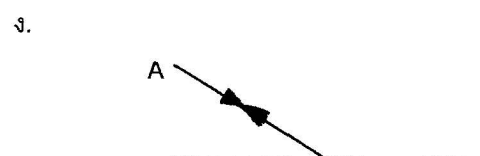
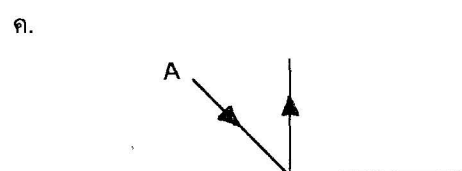
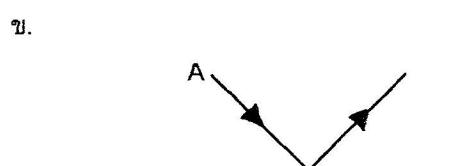
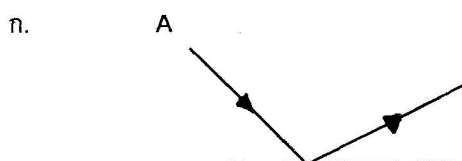
ข้อ	ก	ข	ค	ง
0		X		X

4. ห้ามนักเรียนทำเครื่องหมายใดๆ ลงในแบบทดสอบนี้
5. ให้นักเรียนส่งแบบทดสอบและกระดาษคำตอบคืนกรรมการคุมสอบ
6. ให้นักเรียนทำข้อสอบทุกข้อ เพื่อประโยชน์ของตัวนักเรียนเอง

1. เมื่อใส่น้ำแข็งประมาณ 7 – 8 ก้อนลงในแก้วน้ำทิ้งไว้ 3 นาที ข้อใดไม่ใช่ผลจากการสังเกต
 - ก. มีหยดน้ำเกาะข้างแก้ว
 - ข. น้ำแข็งมีขนาดเล็กลง
 - ค. ปริมาณน้ำมากขึ้น
 - ง. น้ำเปลี่ยนสถานะ
2. ปรากฏการณ์ในข้อใดเกี่ยวข้องกับการเกิดเงามากที่สุด
 - ก. การเกิดรุ้งกินน้ำ
 - ข. การเกิดพระจันทร์ทรงกลด
 - ค. การเกิดฟ้าแลบ ฟ้าร้อง ฟ้าผ่า
 - ง. การเกิดจันทรุปราคา สุริยุปราคา
3. ข้อใดไม่ใช่ผลกระทบจากการใช้ยาฆ่าแมลงที่ได้จากการสังเคราะห์
 - ก. เกิดโรคติดต่อ
 - ข. แหล่งน้ำเป็นพิษ
 - ค. เสียสมดุลธรรมชาติ
 - ง. สารพิษตกค้างในร่างกาย
4. ถ้าน้ำของเหลวชนิดหนึ่งหยดลงบนหินปูน ทำให้เกิดฟองก๊าซ สมมติฐานในข้อใดน่าจะเป็นจริงที่สุด
 - ก. ของเหลวมีสภาพเป็นกรด
 - ข. ของเหลวมีสภาพเป็นเบส
 - ค. ของเหลวมีสภาพเป็นด่าง
 - ง. ของเหลวมีสภาพเป็นกลาง
5. ข้อใดคือผลของการเก็บน้ำอัดลมไว้ในที่ร้อนจัด
 - ก. อาจทำให้รสชาติเสีย
 - ข. อาจทำให้มีการเปลี่ยนสี
 - ค. อาจทำให้เกิดการระเบิดได้
 - ง. อาจทำให้คุณค่าทางอาหารลดลง
6. เมื่อนำน้ำมาต้มจนแห้งไม่มีสิ่งใดอยู่ในภาชนะ ข้อใดคือจุดมุ่งหมายในการทดลองนี้
 - ก. เพื่อทดลองการเปลี่ยนสีของน้ำ
 - ข. เพื่อทดลองการเปลี่ยนสถานะของน้ำ
 - ค. เพื่อทดลองการรวมตัวของน้ำกับความร้อน
 - ง. เพื่อทดลองการเปลี่ยนแปลงของน้ำเป็นสารอื่น
7. เมื่อสสารได้รับความร้อนเพิ่มขึ้น ข้อความใดกล่าวไม่ถูกต้อง
 - ก. สสารจะเปลี่ยนสถานะได้
 - ข. สสารจะเปลี่ยนขนาดได้
 - ค. น้ำหนักของสสารจะเพิ่มขึ้น
 - ง. อุณหภูมิของสสารจะเพิ่มขึ้น
8. เครื่องใช้ไฟฟ้าชนิดใดใช้หลักการเปลี่ยนพลังงานเหมือนกัน
 - ก. ตู้เย็น – เตาหีต
 - ข. เตาอบ – วิทยุ
 - ค. หม้อหุงข้าว – พัดลม
 - ง. เครื่องซักผ้า – เครื่องปั่นน้ำผลไม้
9. เราเรียกสสารที่กั้นขวางทางเดินของแสงว่าอย่างไร
 - ก. ตัวนำแสง
 - ข. ตัวพาแสง
 - ค. ฉนวนแสง
 - ง. ตัวกลางแสง
10. ข้อใดเป็นสาเหตุสำคัญ ที่ทำให้ไฟฟ้าลัดวงจรได้ง่ายที่สุด
 - ก. หลอดไฟขาด
 - ข. เครื่องใช้ไฟฟ้าเก่า
 - ค. เครื่องใช้ไฟฟ้าชำรุด
 - ง. สิ่งห่อหุ้มสายไฟฟ้าชำรุด
11. เมื่อต้มน้ำทิ้งไว้จนเดือด แล้วเพิ่มความร้อนขึ้นอีก จะปรากฏผลอย่างไร
 - ก. จุดเดือดจะต่ำลง
 - ข. จุดเดือดจะสูงขึ้น
 - ค. น้ำจะเปลี่ยนเป็นสารอื่น
 - ง. น้ำจะกลายเป็นไอเร็วขึ้น
12. ปรากฏการณ์ในข้อใดไม่ได้เกิดจากการหักเหของแสง
 - ก. เห็นดินสอหักงอในน้ำ
 - ข. มองภาพตัวเองในกระจกเงาระนาบ
 - ค. มองปลาในน้ำเห็นปลาตื้นกว่าปกติ
 - ง. ปลา มองแมลงบนฝั่ง เห็นแมลงอยู่ไกลออกไป
13. บริเวณผากาต้มน้ำจะมีรูเล็กๆ ไว้เพื่ออะไร
 - ก. สังเกตเวลาน้ำเดือด
 - ข. ลดความดันภายในภา
 - ค. ทำให้เกิดเสียงดังเวลาน้ำเดือด
 - ง. เพิ่มความดันภายในและภายนอกภาให้เท่ากัน
14. สมบัติของสสารด้านใด ที่มีเครื่องวัดได้โดยเฉพาะ
 - ก. มีตัวตน
 - ข. มีน้ำหนัก
 - ค. สัมผัสได้
 - ง. ต้องการที่อยู่

15. ข้อใดไม่เกี่ยวกับความกดดันของอากาศ
- เทน้ำออกจากแก้ว
 - ดูดเศษผงออกจากตู้ปลา
 - การเจาะกระป๋องนมสองรู
 - หยอดตาโดยใช้หลอดหยด
16. ถ้าเราเข้าไปในโรงภาพยนตร์จะมีการเปลี่ยนแปลงของ ม่านตาอย่างไร
- ปิดสนิท
 - ปิดแคบลง
 - เปิดกว้างขึ้น
 - ไม่เปลี่ยนแปลง
17. ข้อใดที่นำกระแสไฟฟ้าจากปฏิกิริยาเคมีมาใช้ประโยชน์
- การใช้พัดลม
 - การใช้ไฟฉาย
 - การใช้หม้อหุงข้าว
 - การใช้เครื่องซักผ้า
18. ถ้าต้องการบันทึกการเปลี่ยนแปลงความกดอากาศ ตลอดวันอย่างต่อเนื่อง ควรใช้เครื่องมือในข้อใด
- बारोगราฟ
 - बारوميเตอร์แบบปรอท
 - बारوميเตอร์แบบตลับ
 - बारوميเตอร์แบบแอนิรอยด์
19. ในการทดลอง “การจัดกราบเบื่อนจากเส้นผ่าด้วย สารเคมี” โดยใช้แอลกอฮอล์ ผงซักฟอก น้ำส้มสายชู เป็นสารเคมีที่ใช้ในการทดลอง ข้อใดต่อไปนี้ ไม่ใช่ ตัวแปรควบคุม
- ระยะเวลาในการแช่สารเคมี
 - ปริมาณของสารเคมีแต่ละชนิด
 - จำนวนครั้งในการซักและการล้าง
 - ขนาดของภาชนะที่ใช้ในการซักล้าง
20. อุณหภูมิในข้อใด ที่ใช้ยึดเป็นหลักในการจำแนกสาร ออกเป็น 3 สถานะ
- ที่อุณหภูมิปกติ
 - ที่ 0 องศาเซลเซียส
 - ต่ำกว่า 0 องศาเซลเซียส
 - ที่อุณหภูมิ 100 องศาเซลเซียส
21. ข้อใด ไม่ใช่ สมบัติทางกายภาพของสสาร
- ขนาด
 - รูปร่าง
 - สถานะ
 - อุณหภูมิ

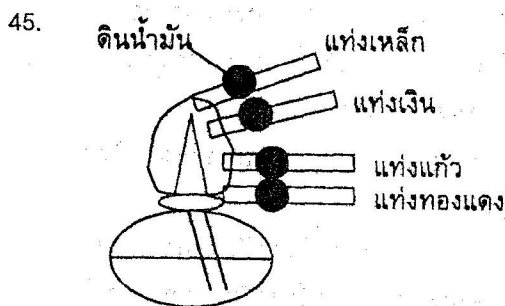
22. อุปกรณ์ไฟฟ้าในข้อใดใช้ตัดกระแสไฟฟ้า เมื่อต้องการ ติดตั้งอุปกรณ์ไฟฟ้าขึ้นใหม่ภายในบ้าน
- ฟิวส์
 - สวิตช์
 - บัลลัสต์
 - สะพานไฟ
23. ภาพในข้อใด แสดงการส่องลำแสงจากตำแหน่ง A ไปที่กระจกเงา



24. แหล่งกำเนิดไฟฟ้าในข้อใดให้กระแสสลับ
- ไดนาโม
 - แบตเตอรี่
 - ปฏิกิริยาเคมี
 - การขัดสีกันของวัตถุ
25. อุปกรณ์ใดที่ใช้วัสดุแตกต่างจากข้ออื่น
- กล้องถ่ายรูป
 - กล้องสลับลาย
 - กล้องจุลทรรศน์
 - กล้องโทรทรรศน์
26. เพราะเหตุใด ในเวลาที่เราดำน้ำจะรู้สึกหุ้อ
- แรงดันของน้ำกดเข้าสู่ร่างกายทุกทิศทาง
 - ความดันอากาศภายนอกสูงกว่าภายในร่างกาย
 - ความดันอากาศภายในสูงกว่าภายนอกในร่างกาย
 - มีน้ำเข้าหูทำให้สภาพการทรงตัวและการได้ยิน เสียไป

27. คนขับรถยนต์ใช้หลักการข้อใด ในการปรับกระจกมองหลัง
- ให้กระจกปรับแสงตกกระทบจากคนขับ
 - ให้กระจกหันไปทางคนขับ เพื่อเห็นภาพในถนน
 - ให้กระจกหันมาทางคนขับ สะท้อนภาพไปในถนน
 - ให้กระจกปรับแสงตกกระทบสะท้อนมาที่คนขับ
28. ภาพของวัตถุที่ไปปรากฏที่เรตินามีลักษณะอย่างไร
- ภาพจริง หัวตั้ง
 - ภาพจริง หัวกลับ
 - ภาพเสมือน หัวตั้ง
 - ภาพเสมือน หัวกลับ
29. เพราะเหตุใดเราจึงไม่รู้สึกว่ามีอากาศกดดันอยู่รอบตัว
- เพราะอากาศภายนอกร่างกายมีแรงดันแต่ภายในไม่มี
 - เพราะอากาศภายในร่างกายมีแรงดันเท่ากับอากาศภายนอก
 - เพราะอากาศภายในร่างกายมีแรงดันน้อยกว่าอากาศภายนอก
 - เพราะอากาศภายนอกในร่างกายมีแรงดันน้อยกว่าอากาศภายใน
30. วัตถุสีใดที่มีอัตราการแผ่รังสีความร้อนมากที่สุด
- ดำ
 - แดง
 - ขาว
 - เขียว
31. เมื่ออากาศร้อนจะทำให้ความกดอากาศเป็นอย่างไร
- ลดลง
 - เท่าเดิม
 - เพิ่มขึ้น
 - ไม่แน่นอน
32. คำโฆษณาว่าฆ่าแมลงข้อใดที่ไม่เป็นความจริง
- ฆ่าแมลงสาบได้ 100%
 - มีกลิ่นหอมเหมือนดอกไม้
 - ออกฤทธิ์ทันทานนับสัปดาห์
 - ไม่เป็นอันตรายต่อคนและสัตว์เลี้ยง
33. การผลิตกระแสไฟฟ้าโดยอาศัยแรงดันไอน้ำส่งผลต่อสิ่งใดมากที่สุด
- มลพิษทางน้ำ
 - มลพิษทางเสียง
 - มลพิษทางอากาศ
 - มลพิษทางทัศนียภาพ
34. การผลิตกระแสไฟฟ้าระบบใด ทำให้เกิดมลพิษต่อสิ่งแวดล้อมน้อยที่สุด
- ระบบพลังน้ำ
 - ระบบพลังไอน้ำ
 - ระบบกังหันก๊าซ
 - ระบบพลังนิวเคลียร์
35. น้ำที่เหลือจากการซักผ้ามีประโยชน์ต่อพืชอย่างไร
- ทำให้ดินร่วน
 - ช่วยฆ่าวัชพืช
 - ทำให้ดินมีฟองอากาศ
 - ช่วยให้พืชเจริญเติบโตได้ดี
36. สายล่อฟ้าช่วยลดอันตรายจากปรากฏการณ์ฟ้าผ่าได้อย่างไร
- ป้องกันไม่ให้ไฟฟ้าจากธรรมชาติไหลเข้าสู่สายไฟ
 - ช่วยให้ไฟฟ้าจากธรรมชาติสะท้อนกลับสู่บรรยากาศ
 - ช่วยเปลี่ยนแรงดันไฟฟ้าจากธรรมชาติก่อนเข้าสู่บ้านเรือน
 - สามารถควบคุมปริมาณกระแสไฟฟ้าไม่ให้เข้าสู่เครื่องใช้ไฟฟ้า
37. วิธีการใดต่อไปนี้ไม่ช่วยให้หายจากอาการหูดื้อ ในขณะที่เดินทางบนภูเขาสูง
- การหา
 - การทานขนม
 - การกลั้วหายใจ
 - การเคี้ยวหมากฝรั่ง
38. ข้อใดที่แสดงว่าสารต้องการที่อยู่
- น้ำผสมกับน้ำตาลเป็นน้ำเชื่อม
 - เมื่อเป่าลมเข้าไปในลูกโป่งจะพองออก
 - ของเหลวที่มีปริมาตรคงที่ในภาชนะใดๆ
 - ใส่ก้อนหินลงในแก้วน้ำให้ระดับน้ำสูงขึ้น
39. การที่เราดูดน้ำหวานขึ้นไปตามหลอดดูดได้ - ด้วยเหตุผลในข้อใด
- อากาศภายในและอากาศภายนอกหลอดมีแรงดันเท่ากัน
 - อากาศภายนอกหลอดมีแรงดันน้อยกว่าภายใน
 - อากาศภายนอกหลอดมีแรงดันมากกว่าภายใน
 - อากาศภายนอกหลอดได้รับความกดอากาศสูงจากบริเวณอื่น

40. ลักษณะใดของบ้านทรงไทยที่เป็นไปตามการถ่ายเทความร้อน
- ใต้ถุนสูง
 - เพดานสูง
 - นอกชานกว้าง
 - สร้างด้วยไม้สัก
41. โซฟอน คืออะไร
- การถ่ายเทก๊าซจากที่ต่ำไปที่สูง
 - การถ่ายเทก๊าซจากที่สูงไปที่ต่ำ
 - การถ่ายเทของเหลวจากที่ต่ำไปที่สูง
 - การถ่ายเทของเหลวจากที่สูงไปที่ต่ำ
42. เมื่อใดที่แสงเดินทางเป็นเส้นตรง
- เมื่อผ่านตัวกลางที่มีสมบัติต่างกัน
 - เมื่อผ่านตัวกลางที่มีสมบัติเหมือนกัน
 - เมื่อออกจากแหล่งกำเนิดสู่สายตาคน
 - เมื่อออกจากแหล่งกำเนิดผ่านตัวกลางโปร่งแสง
43. เราใช้ประโยชน์จากการหักเหของแสงในเรื่องใด - มากที่สุด
- การทำแว่นตา
 - การส่งสัญญาณ
 - การแทงปลาในน้ำ
 - การสร้างภาพลวงตา
44. การส่งผ่านความร้อนแบบใด ที่โมเลกุลของตัวกลางมีการเคลื่อนที่
- การนำความร้อน
 - การพาความร้อน
 - การแผ่รังสีความร้อน
 - การนำและการพาความร้อน



จากรูป เป็นการทดลองเรื่อง “การเปรียบเทียบการนำความร้อนของสสาร” สิ่งใดที่ไม่ต้องควบคุมให้เท่ากัน

- ความยาวของแท่งวัตถุ
- ระยะห่างของก้อนดินน้ำมัน
- ปริมาณแอลกอฮอล์ในตะเกียง
- การลุกไหม้ของไฟจากตะเกียงแอลกอฮอล์

46. ข้อใดไม่ใช่ผลการสังเกตที่ได้จากการอ่านค่าของ บารอมิเตอร์
- ความกดอากาศ
 - ลักษณะอากาศ
 - อุณหภูมิของอากาศ
 - ระดับความสูงของพื้นที่
47. ถ้าแก้วสองใบซ้อนติดกันดังรูป แล้วดึงไม่ออก จะต้องทำอย่างไร



- นำแก้วทั้งสองใบแช่ในอ่างน้ำเย็น
 - นำแก้วทั้งสองใบไปต้มในอ่างน้ำร้อน
 - นำน้ำแข็งใส่แก้วใบในแล้วแช่ในอ่างน้ำอุ่น
 - นำน้ำอุ่นใส่แก้วใบในแล้วแช่ในอ่างน้ำเย็น
48. ข้อใดเป็นสมบัติทางเคมีของสสาร
- สถานะ
 - จุดเดือด
 - การเป็นสนิม
 - การละลายในน้ำ
49. ข้อใดเป็นการเปลี่ยนแปลงทางกายภาพ
- เหล็กขึ้นสนิม
 - การขูดมะพร้าว
 - ลูกเหม็นระเหิด
 - รอยเกล็ดลงในเปลวไฟ
50. ถ้านำไฟฉายส่องผ่านกระดาษลอกลาย จะปรากฏผลอย่างไร
- เกิดการหักเหของแสง
 - เกิดการกระจายแสง
 - เกิดการกระเจิงแสง
 - เกิดการสะท้อนกลับ

ประวัติย่อผู้วิจัย

ประวัติผู้วิจัย

ชื่อ นางสาวกัญญรัตน์ ทรงประโคน

เกิดวันที่ 21 เดือนตุลาคม พุทธศักราช 2517

สถานที่เกิด อำเภอเมือง จังหวัดชัยภูมิ

สถานที่อยู่ปัจจุบัน 70/1 ซ.ราชวิถี 40 ถ.ราชวิถี

แขวงวชิระ เขตดุสิต

กรุงเทพฯ 10300

สถานที่ทำงานในปัจจุบัน โรงเรียนราชินี 444 ถ.มหาราช แขวงพระบรมมหาราชวัง

เขตพระนคร กรุงเทพฯ 10200

ประวัติการศึกษา

พ.ศ. 2532	มัธยมศึกษาปีที่ 3 โรงเรียนสตรีชัยภูมิ
พ.ศ. 2536	มัธยมศึกษาปีที่ 6 โรงเรียนชัยภูมิภักดีชุมพล
พ.ศ. 2540	ค.บ.ประถมศึกษา (เกียรตินิยมอันดับ 2) สถาบันราชภัฏสวนดุสิต
พ.ศ. 2545	กศ.ม.(การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ