

การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลิ่งที่แตกต่างกัน 3 แบบ

ปริญญาณิพนธ์
ของ
พิมพ์ร ภูมิรา

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

เมษายน 2544

ลิขสิทธิ์เป็นของ มหาวิทยาลัยศรีนครินทรวิโรฒ

342.9044

พ 414 ก

ร. 3

4

การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกัน 3 แบบ

27 ก.ย. 2544

บทคัดย่อ
ของ
พิมพ์ร ภูมิรา

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
เมษายน 2544

พิมพร ภูมิรา. (2544). การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
 วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกัน 3
 แบบ. ปริญญา นิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย
 มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์ ดร.บุญเชิด
 ภิญโญนนตพงษ์, รองศาสตราจารย์นิภา ศรีไพโรจน์

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์
 ทางการเรียนวิชาคณิตศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกัน 3 แบบ คือ 1 กฎชิลลิ่ง
 แบบคะแนนศูนย์ต่อเนื่อง 2 กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และ
 3 กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่
 สอง ตำแหน่งที่สามและการไม่กำหนดตำแหน่งการตอบผิด กลุ่มตัวอย่างที่ใช้ในการศึกษา
 ได้แก่นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงาน
 การประถมศึกษาจังหวัดกาญจนบุรี จำนวน 790 คนได้มาโดยการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้
 เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ผลการวิจัยสรุปได้ดังนี้

1. ความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา
 คณิตศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบที่ 1 จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่
 สอง ตำแหน่งที่สามและการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทาง
 สถิติที่ระดับ .05 โดยความเชื่อมั่นของการตอบผิดตำแหน่งที่หนึ่งมีค่าสูงสุดเท่ากับ .96
2. ความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา
 คณิตศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบที่ 2 จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่
 สอง ตำแหน่งที่สามและการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทาง
 สถิติที่ระดับ .05 โดยความเชื่อมั่นของการตอบผิดตำแหน่งที่หนึ่งมีค่าสูงสุดเท่ากับ .96
3. ความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา
 คณิตศาสตร์ที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบที่ 3 จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่
 สอง ตำแหน่งที่สามและการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทาง
 สถิติที่ระดับ .05 โดยความเชื่อมั่นของการตอบผิดตำแหน่งที่หนึ่งมีค่าสูงสุดเท่ากับ .95
4. ความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันทั้ง 3 แบบ
 ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์จากการตอบผิดตำแหน่งที่หนึ่งมีค่า
 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
5. ความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่ต่างกันทั้ง 3 แบบ
 ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์จากการตอบผิดตำแหน่งที่สองมีค่า
 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

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

AN ABSTRACT
BY
PIMPORN PUMIRA

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

April 2001

Pimporn Pumira. (2001). *The Effects of Different Ceiling Rules on the Internal Consistency Reliability of Mathematics 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 the research was to study the effects on the reliability of three types of 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 one single failure found within 3 time consecutively. Type 3, the ceiling rule determined according to more than one consecutive failure found within 3 time consecutively. The unlimited failure or no ceiling was also considered as a control group within three types of different ceiling rules. The sample was 790 Prathom Suksa VI students in Kanchanaburi Province selected by stratified random sampling technique. The instrument used in this study was a mathematics achievement test.

The results were as follows :

1. The reliability of mathematics achievement test scored by using ceiling rule type 1 was significant difference at .05 level. The first consecutive failure found yield the highest reliability at .96.
2. The reliability of mathematics achievement test scored by using ceiling rule type 2 was significant difference at .05 level. The first single failure found yield the highest reliability at .96.
3. The reliability of mathematics achievement test scored by using ceiling rule type 3 was significant difference at .05 level. The first time of more than one consecutive failure found yield the highest reliability at .95.
4. The reliability of mathematics achievement test of three different ceiling rules that using the one consecutive failure found - type 1, the first time of single failure found - type 2, and the first time of more than one consecutive failure found - type 3 was significant difference at .05 level.

5. The reliability of mathematics achievement test of three different ceiling rules that using the two consecutive failure found - type 1, the second time of single failure found - type 2, and the second time of more than one consecutive failure found - type 3 was significant difference at .05 level.

6. The reliability of mathematics achievement test of three differences ceiling rules that using the three consecutive failure - type 1, the third time of single failure found - type 2, and the third time of more than one consecutive failure found - type 3 was significant difference at .05 level.

ปริญญานิพนธ์

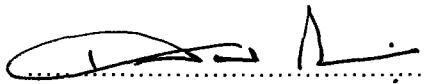
เรื่อง

การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน 3 แบบ

ของ

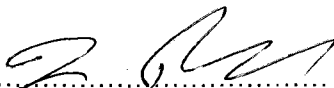
นางพิมพ์ ภูมิรา

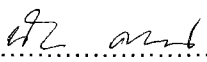
ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

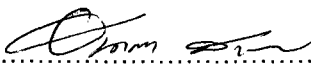
 คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

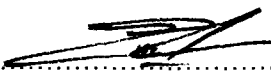
วันที่ 30 เมษายน พ.ศ. 2544

คณะกรรมการสอบปริญญานิพนธ์

 ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์)

 กรรมการ
(รองศาสตราจารย์ นิภา ศรีไพโรจน์)

 กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ อังคณา สายยศ)

 กรรมการที่แต่งตั้งเพิ่มเติม
(ผู้ช่วยศาสตราจารย์ ดร. นิคม ตั้งกะพิภพ)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงลงได้ด้วยดี เพราะการได้รับความอนุเคราะห์ความรู้ในด้านต่าง ๆ จาก รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนันตพงษ์ รองศาสตราจารย์ นิภา ศรีไพโรจน์ ตลอดจนคณาจารย์ภาควิชาการวัดผลและวิจัยการศึกษาทุกท่านที่ได้กรุณาให้ความคิดเห็น ข้อเสนอแนะ การแก้ไขข้อบกพร่อง ซึ่งผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอกราบขอบพระคุณรองศาสตราจารย์ อังคณา สายยศ และผู้ช่วยศาสตราจารย์ ดร.นิคม ดังคะพิภพ กรรมการสอบปากเปล่าที่ให้แนวความคิดสร้างสรรค์ และชี้แนะข้อบกพร่อง

ขอกราบขอบพระคุณ ดร.สุพร เข้มเฮง ดร.สมบัติ สุวรรณพิทักษ์ อาจารย์ชวลิต รวยอาจิน อาจารย์ขรรค์ชัย คงเสน่ห์ และอาจารย์ธานี เขียนมัน ที่ได้กรุณาเป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือการวิจัยพร้อมทั้งแสดงข้อคิดเห็น พร้อมทั้ง ข้อเสนอแนะที่เป็นประโยชน์ในการปรับปรุงแก้ไขเครื่องมือให้มีประสิทธิภาพยิ่งขึ้น

กราบขอบพระคุณ รองศาสตราจารย์ ดร.สมบุรณ์ สุริยวงศ์ ที่ให้ความรู้เกี่ยวกับวิธีการตรวจให้คะแนนโดยกฎซิกลิ่ง ตลอดจนการวิเคราะห์ข้อมูลในโปรแกรม SPSS / PC⁺

ขอขอบพระคุณ รองศาสตราจารย์ ดร.ยงยุทธ ดันจุลเวสส พร้อมทั้ง อาจารย์สุชาติน้ำประสานไทย ที่ได้กรุณาช่วยตรวจสอบความถูกต้องการจัดพิมพ์ การใช้ภาษา และช่วยเหลือเป็นกำลังใจแก่ผู้วิจัยด้วยดี ตลอดมา

ขอขอบคุณผู้บริหารโรงเรียน คณะครูผู้สอนในโรงเรียนที่เป็นกลุ่มตัวอย่างและนักเรียนในสังกัด สำนักงานการประถมศึกษาจังหวัดกาญจนบุรี ที่ให้การสนับสนุนและร่วมมือช่วยเหลือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

สารประโยชน์จากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบให้แก่ครู อาจารย์ และนิสิต นักศึกษาทุกท่าน ทุกคน ที่สนใจศึกษา เพื่อนำไปประกอบการเรียนรู้ และสุดท้ายผู้วิจัยขออุทิศคุณความดีนี้แด่คุณแม่ซึ่งเปรียบเสมือนครูคนแรก และคุณพ่อผู้ล่วงลับ

พิมพ์ร ภูมิรา

สารบัญ

บทที่		หน้า
1	บทนำ.....	1
	ภูมิหลัง.....	1
	ความมุ่งหมายของการศึกษาค้นคว้า.....	5
	ความสำคัญของการศึกษาค้นคว้า.....	6
	ขอบเขตของการศึกษาค้นคว้า.....	6
	ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย.....	6
	ตัวแปรที่ใช้ในการศึกษาค้นคว้า.....	6
	นิยามศัพท์เฉพาะ.....	7
2	เอกสารและงานวิจัยที่เกี่ยวข้อง.....	11
	วิธีการตรวจให้คะแนน.....	11
	การเรียงข้อคำถามในแบบทดสอบ.....	22
	การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้.....	24
	ความเชื่อมั่นของแบบทดสอบ.....	36
	ทฤษฎีการตอบข้อคำถาม.....	55
	งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนน.....	59
	งานวิจัยที่เกี่ยวข้องกับการจัดเรียงข้อคำถาม.....	68
	สมมติฐานในการศึกษาค้นคว้า.....	75
3	วิธีดำเนินการศึกษาค้นคว้า.....	77
	ประชากร.....	77
	กลุ่มตัวอย่าง.....	77
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	79
	วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	80
	วิธีดำเนินการเก็บรวบรวมข้อมูล.....	88
	วิธีการตรวจให้คะแนนโดยกฎซึลลิ่ง 3 แบบ.....	89
	ขั้นตอนในการวิเคราะห์ข้อมูล.....	95
	สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	96

สารบัญ(ต่อ)

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล.....	97
สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล	97
ผลการวิเคราะห์ข้อมูล.....	97
5 สรุป อภิปรายผล และข้อเสนอแนะ.....	113
ความมุ่งหมาย สมมติฐาน และเครื่องมือที่ใช้ในการเก็บข้อมูล	113
สรุปผลการศึกษาค้นคว้า	115
อภิปรายผล.....	118
ข้อเสนอแนะ.....	122
บรรณานุกรม.....	123
ภาคผนวก	133
ภาคผนวก ก.....	134
รายนามผู้เชี่ยวชาญประเมินความสอดคล้องระหว่างข้อคำถามกับ	
พฤติกรรมการเรียนรู้.....	135
ภาคผนวก ข.....	136
ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้	
(INDEX OF CONSISTENCY : IOC).....	137
ภาคผนวก ค.....	138
ค่าความยาก (ITEM DIFFICULTY : B) ของแบบทดสอบวัดผลสัมฤทธิ์ทาง	
การเรียนรู้วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่วิเคราะห์โดย	
ราชคิโมเดล (RASCH MODEL) จากการทดลองใช้ (TRY OUT).....	139
ภาคผนวก ง.....	141
แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษา	
ปีที่ 6	142
ภาคผนวก จ.....	160
หนังสือเชิญเป็นผู้เชี่ยวชาญ	161
ภาคผนวก ฉ.....	165
หนังสือขอความร่วมมือเพื่อการวิจัย	166

สารบัญ(ต่อ)

บทที่	หน้า
ประวัติย่อผู้วิจัย.....	167

บัญชีตาราง

ตาราง	หน้า
1 วิธีการให้คะแนนแบบเลือกเสรี	12
2 วิธีให้คะแนนแบบบอกระดับความมั่นใจ	12
3 วิธีให้คะแนนแบบกิบบอนส์ โอลส์คิน และโซเบล	14
4 วิธีการตรวจให้คะแนนของเกริกชัย ชวบเจริญ	19
5 วิธีการตรวจให้คะแนนของสำราญ มีแจ้ง	19
6 จำนวนนักเรียนกลุ่มตัวอย่างที่ใช้ในการทดสอบ จำแนกตามขนาดของโรงเรียน	78
7 ตารางแสดงการวิเคราะห์รายละเอียดวิชาคณิตศาสตร์	82
8 ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน 3 แบบ ที่กำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนด ตำแหน่งการตอบผิด	99
9 ค่าความเชื่อมั่น และค่าที่แปลงเป็นคะแนนมาตรฐานซีของแบบทดสอบที่ตรวจ ให้คะแนนโดยกฎซิลลิงที่แตกต่างกัน 3 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด	101
10 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้ คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิด ต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด	103
11 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดย กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด	104
12 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้ คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่ง การตอบผิด	105
13 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดย กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด	106

บัญชีตาราง(ต่อ)

ตาราง		หน้า
14	การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด	107
15	การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด	108
16	การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง	109
17	การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง	109
18	การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สอง	110
19	การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สอง	111
20	การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สาม	111
21	การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สาม	112

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 ภาพแสดงความแตกต่างระหว่างความสามารถของคนกับความยากของ ข้อสอบ	57
2 ลำดับขั้นตอนในการดำเนินการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เนื้อหาบทที่ 1 - 8 ตามพฤติกรรม การเรียนรู้ของวิลสัน	80

บทที่ 1

บทนำ

ภูมิหลัง

การพัฒนาคุณภาพมนุษย์เป็นสิ่งสำคัญที่สุด (สำนักงานคณะกรรมการการศึกษาแห่งชาติ. 2540 : 24) เพราะประเทศชาติจะพัฒนาได้ต้องอาศัยคนดีมีคุณภาพ และการที่คนจะดีมีคุณภาพได้ก็ต้องอาศัยการศึกษา (จิตรกร ตั้งเกษมสุข. 2525 : 201) ส่วนการศึกษาจะประสบความสำเร็จได้มากน้อยเพียงใดต้องอาศัยการวัดและการประเมินผล เพราะการวัดและการประเมินผลมีเพื่อพัฒนาสมรรถภาพของมนุษย์ (ชวาล แพรัตกุล. 2516 : 3) เกี่ยวกับพฤติกรรมด้านการเรียนรู้ด้วยการวัดหลาย ๆ ครั้งหลาย ๆ วิธี เพื่อช่วยให้การคาดคะเนหรือการแปลความก้าวหน้าของผู้เรียนว่าบรรลุผลตรงตามเป้าหมายที่แตกต่างกันออกไปได้ นอกจากนี้ยังเป็นการวินิจฉัยคุณภาพกระบวนการเรียน การสอนในแต่ละขั้นตอนว่ามีประสิทธิภาพเพียงใด การวัดและการประเมินผลด้านพฤติกรรมศาสตร์ และสังคมศาสตร์เป็นการวัดผลทางอ้อม ย่อมเกิดความคลาดเคลื่อนได้ง่าย (ล้วน สายยศ และอังคณา สายยศ. 2539 : 11) เครื่องมือวัดผลที่นิยมใช้กันมากก็คือ แบบทดสอบ (Tyler. 1971 : 26) ดังนั้น จึงจำเป็นอย่างยิ่งที่แบบทดสอบที่จะนำมาใช้ในการวัดผลต้องมีคุณภาพเป็นที่เชื่อถือได้ว่า ให้คะแนนที่เป็นความสามารถที่แท้จริงได้มากโดยมีความคลาดเคลื่อนในการวัดน้อยที่สุด (อวยพร วิบูลย์กาญจน์. 2526 : 7) แบบทดสอบปรนัยแบบเลือกตอบ เป็นแบบทดสอบที่ดีที่สุดในแบบหนึ่ง (ชวาล แพรัตกุล. 2518 : 164) ถ้าผู้สร้างแบบทดสอบมีความรู้และมีความชำนาญ ก็สามารถสร้างให้วัดสมรรถภาพทางสมองด้านต่าง ๆ ตามลำดับความรู้ได้อย่างเหมาะสม ทั้งยังได้ผลการทดสอบที่มีความเที่ยงตรง ความเชื่อมั่นสูง เนื่องจากวัดได้ตรงตามเนื้อหา ตรงตามจุดประสงค์การเรียนรู้ สะดวกในการตรวจให้คะแนน (สมพงษ์ อิศวฤตฤกษ์. 2542 : 2) ใช้ได้กับนักเรียนทุกระดับ เหมาะกับการสอบที่มีผู้สอบเป็นจำนวนมาก (Ebel. 1965 : 60) และให้ความยุติธรรมแก่ผู้สอบ (ชวาล แพรัตกุล. 2516 : 150)

แม้ว่าแบบทดสอบปรนัยแบบเลือกตอบจะมีข้อดีอยู่หลายประการ แต่นักการศึกษาหลายท่านยังเห็นว่ามีข้อเสียด้านส่งเสริมการเดา (พจน์ สะเพียรชัย. 2515 : 21) ทำให้ไม่สามารถแยกผู้ที่มีความรู้ได้ถูกต้อง ลิทเทิล (วัชรารณ จิตรมาศ. 2538 : 2 ; อ้างอิงจาก Little 1962) กล่าวถึงลักษณะระดับความรู้ของผู้ตอบที่นำไปสู่การตอบถูกหรือผิดเกี่ยวข้องกัน 6 ลักษณะ คือ มีความรู้ถูกต้องในเรื่องนั้นนำไปสู่การตอบที่ถูก มีความรู้บางส่วนนำไปสู่การตอบที่ถูก มีความรู้บางส่วนนำไปสู่การตอบที่ผิด ไม่มีความรู้ทำให้เลือกตอบแบบเดาสุ่ม มีความรู้ไม่ถูกต้องนำไปสู่การตอบผิด รู้สึกว่าคำถามในข้อสอบไม่เหมาะสม รู้คำตอบที่ผู้วัดต้องการให้ตอบแต่ไม่เห็นด้วยจึงไม่เลือกคำตอบที่ถูกต้อง นอกจากนี้ในการตอบข้อสอบนักเรียนจะพิจารณาหาข้อที่ถูกที่

สุดเพียงข้อเดียวเมื่อพบแล้วจะไม่พิจารณาตัวเลือกอื่น ๆ โดยจะถือว่าตัวเลือกที่เหลือผิดทันที ถ้านักเรียนไม่ทราบมาก่อนว่าทำไมจึงผิด หรืออาจตอบโดยค่อย ๆ กำจัดตัวเลือกที่ผิดออกไป ซึ่งทำได้ง่ายเพราะโอกาสที่เหลือตัวเลือกตอบต่าง ๆ เป็นตัวเลือกที่ผิดนั้นมีมาก ทำให้ผลจากการวัด เป็นคะแนนจริงรวมกับคะแนนความคลาดเคลื่อนไม่ใช่ความสามารถที่แท้จริงของผู้เรียน จึงต้องมีการคิดวิธีแก้การเดาหลายรูปแบบ เช่น สร้างข้อสอบให้มีตัวเลือกเป็นเอกพันธ์ สร้างจำนวนตัวเลือกหลาย ๆ ตัว หรือใช้สูตรแก้การเดา

ในการทดสอบแต่ละครั้ง คะแนนที่นักเรียนสอบได้ (X) จะประกอบด้วยคะแนนจริง (T) และคะแนนความคลาดเคลื่อน (E) ดังสมการ $X = T + E$ (Gulliksen. 1950 ; Allen and Yen.1979.) คะแนนที่นักเรียนสอบได้และคะแนนความคลาดเคลื่อนเป็นค่าที่แปรเปลี่ยนไปตามบุคคลในการสอบแต่ละครั้ง คะแนนจริงและคะแนนความคลาดเคลื่อนไม่มีความสัมพันธ์กันทั้งยัง ไม่สัมพันธ์กับคะแนนจริงและคะแนนความคลาดเคลื่อนในการทดสอบครั้งอื่น ๆ ดังนั้น คะแนนจริงของบุคคลจะเกิดขึ้นได้ก็ต่อเมื่อไม่มีการคลาดเคลื่อนในการวัด จากทฤษฎีคะแนนจริง มาตรฐานเดิมที่ว่าความเชื่อมั่นของแบบทดสอบ คือ สัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สอบได้ สัดส่วนนี้จะมีค่ามากหรือน้อยขึ้นอยู่กับความแปรปรวนของคะแนนความคลาดเคลื่อน นั่นคือ เมื่อความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นความแปรปรวนของคะแนนความคลาดเคลื่อนก็จะมีจำนวนน้อยลง ในทางตรงกันข้ามถ้าความแปรปรวนของคะแนนความคลาดเคลื่อนมากขึ้น ค่าสัมประสิทธิ์ของความเชื่อมั่นของแบบทดสอบก็จะลดลง

ปัญหาเรื่องการเดามีผลต่อค่าความเชื่อมั่นและคะแนนความคลาดเคลื่อน ถ้าสามารถกำหนดเงื่อนไขในการให้นักเรียนตอบและเงื่อนไขการตรวจให้คะแนน ได้อย่างถูกต้องเหมาะสม จะช่วยให้คะแนนที่ได้มีค่าสหสัมพันธ์สูงกับความสามารถที่แท้จริงของนักเรียน (อรรพรรณ ตัณฑิ เจริญรัตน์. 2517 : 2 - 3) สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) วิลสันและคณะ (Wilson and others. 1988 : 213 - 217) เสนอวิธีการตรวจให้คะแนนโดยกฎซิลลิง (Ceiling Rules) ที่เรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดตามรูปแบบของกัตแมน (Guttman Model) โดยถือว่าคำตอบข้อที่นักเรียนตอบถูกต้องที่อยู่หลังคำตอบข้อที่นักเรียนตอบผิดเป็นการตอบถูกต้องด้วยการเดามากกว่าเป็นการตอบจากความสามารถที่แท้จริง การตอบผิดของแต่ละบุคคลจะมีระดับต่าง ๆ กัน ขึ้นอยู่กับความสามารถของผู้ตอบและระดับความยากของแบบทดสอบที่ใช้

มีงานวิจัยในต่างประเทศเกี่ยวกับวิธีการตรวจให้คะแนนโดยกฎซิลลิงสมบูรณ์ 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 จะไม่นำมานับคะแนน สำหรับการไม่กำหนดตำแหน่งการตอบผิดจะให้ข้อละ 1 คะแนนกับทุกข้อที่ตอบถูก ผลที่ได้จากการศึกษาวิจัยพบว่า การตอบผิดตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่ง และการไม่กำหนดตำแหน่งการตอบผิดโดยการตอบผิดตำแหน่งที่สองให้ค่าความเชื่อมั่นสูงสุด

นอกจากนี้วิลสัน และคณะ (Wilson and others. 1988 : 213 - 216) ได้ศึกษาวิจัยค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์กับนักเรียนเกรด 8 โดยเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด แล้วตรวจให้คะแนนโดยกฎชิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด เพื่อหาความสามารถที่แท้จริงโดยให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด (0) เป็นข้อแรกจึงหยุดตรวจแล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่หนึ่ง คำตอบถูกหลังการตอบผิดตำแหน่งที่หนึ่งจะไม่นำมานับคะแนน และให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรกจึงหยุดตรวจ แล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่สอง คำตอบถูกหลังการตอบผิดตำแหน่งที่สองจะไม่นำมานับคะแนน ในทำนองเดียวกันการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่และตำแหน่งที่ห้าจะให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิดติดต่อกันตั้งแต่สามข้อ (000) สี่ข้อ (0000) และห้าข้อ (00000) เป็นครั้งแรกจึงหยุดตรวจ แล้วนับคะแนนโดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าตามลำดับ คำตอบถูกหลังการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าจะไม่นำมานับคะแนน สำหรับการไม่กำหนดตำแหน่งการตอบผิดจะให้ข้อละ 1 คะแนนกับทุกข้อที่ตอบถูก ผลการศึกษาวิจัยที่ได้สอดคล้องกับ สมบูรณ์ สุริยวงศ์ คือ การตอบผิดตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ และ

ตำแหน่งที่ห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่ง และการไม่กำหนดตำแหน่งการตอบผิด โดยการตอบผิดตำแหน่งที่สองให้ค่าความเชื่อมั่นสูงสุด

สำหรับงานวิจัยในประเทศไทยเกี่ยวกับวิธีการตรวจให้คะแนนโดยกฎซิลลิงมีผู้ศึกษาวิจัยไว้ 3 ท่าน คือ จินตนา ต่อทรัพย์สินชัย (2533) อัมพร ยิ้มแย้ม (2533) และมานะ คิวเจริญ (2539) ซึ่งทั้ง 3 ท่านได้ศึกษาวิธีการตรวจให้คะแนนโดยกฎซิลลิงของ สมบูรณ์ สุริยวงศ์ วิลสัน และ คณะ (กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง) แล้วนำมาศึกษาวิจัยกับนักเรียนระดับชั้นมัธยมศึกษา โดยเรียงข้อสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์จากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดและกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เพื่อหาความสามารถที่แท้จริงของนักเรียน และให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด (0) เป็นข้อแรกจึงหยุดตรวจแล้วนับคะแนน โดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่หนึ่ง คำตอบถูกหลังการตอบผิดตำแหน่งที่หนึ่งจะไม่นำมานับคะแนน และให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด (0,0) เป็นข้อที่สอง จึงหยุดตรวจแล้วนับคะแนน โดยถือว่าข้อที่ตอบผิดนั้นเป็นการตอบผิดตำแหน่งที่สองคำตอบถูกหลังการตอบผิดตำแหน่งที่สองจะไม่นำมานับคะแนน ในทำนองเดียวกันตำแหน่งการตอบผิดตำแหน่งที่สามตำแหน่งที่สี่ และตำแหน่งที่ห้าจะให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 ถึงข้อที่นักเรียนตอบผิด (0,0,0,0) เป็นข้อที่สาม (0,0,0,0) ข้อที่สี่ และ (0,0,0,0,0) ข้อที่ห้าจึงหยุดตรวจ แล้วนับคะแนน โดยถือว่าข้อที่ตอบผิดนั้น ๆ เป็นการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าตามลำดับ คำตอบถูกหลังการตอบผิดตำแหน่งที่สาม ตำแหน่งที่สี่ และตำแหน่งที่ห้าจะไม่นำมานับคะแนน ส่วนการไม่กำหนดตำแหน่งการตอบผิดจะให้ข้อละ 1 คะแนนกับทุกข้อที่ตอบถูก ผลจากการศึกษาวิจัยที่ได้ของทั้ง 3 ท่านแตกต่างกันออกไป ซึ่งจะเห็นได้ว่าวิธีการตรวจให้คะแนนโดยกฎซิลลิงในประเทศแตกต่างจากของต่างประเทศและขาดเหตุผลรองรับในที่นี้ ผู้วิจัยจึงขอเรียกการตรวจให้คะแนนแบบนี้ว่ากฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง

ในปัจจุบันวิธีการตรวจให้คะแนนแบบทดสอบที่ใช้กันอยู่เป็นการให้คะแนนแบบ 0 - 1 คือ ให้ 1 คะแนนสำหรับคำตอบถูก และ 0 คะแนนสำหรับคำตอบผิด ไม่มีการกำหนดตำแหน่งการตอบผิดผู้วิจัยจึงมีความเห็นว่าวิธีการตรวจให้คะแนนเช่นนี้ เป็นวิธีการที่น่าจะมีความคลาดเคลื่อนมากเพราะเป็นการปฏิเสธคะแนนที่เป็นจริง (Type I error) และยอมรับคะแนนที่เป็นเท็จ (Type II error) เข้ามาเป็นคะแนนรวม อย่างไรก็ตาม วิธีการตรวจให้คะแนนโดยกฎซิลลิงของ สมบูรณ์ สุริยวงศ์ วิลสัน และคณะ ซึ่งค่อนข้างจะเป็นวิธีการที่เข้มงวดมากเกินไปจึงเป็นการให้โอกาสการตอบผิดที่ต่อเนื่องกันเพียงครั้งเดียวเป็นจุดตัด ผู้วิจัยเห็นว่าน่าจะเปิดโอกาสการตอบผิดที่ต่อเนื่องกันหลาย ๆ ครั้งเป็นจุดตัด เพื่อหาความสามารถที่เป็นคะแนนจริงได้มากขึ้น จากแนวคิดดังกล่าว ผู้วิจัยจึงมีความคิดที่จะศึกษาค่าความเชื่อมั่นของแบบทดสอบ

จากผลการใช้กฎซึลลิ่ง ของ สมบูรณ์ สุริยวงศ์ วิลสัน และ คณะ (กฎซึลลิ่งแบบคะแนนศูนย์ ต่อเนื่อง) กฎซึลลิ่งที่ผู้วิจัยในประเทศไทยได้ศึกษาไว้ (กฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียง ตัวเดียวหลาย ๆ ครั้ง) กับกฎที่เปิดโอกาสการตอบผิดที่ต่อเนื่องกันได้มากกว่า 1 ครั้งติดต่อกัน (กฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว) ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) จะให้ค่าความเชื่อมั่นต่างกันหรือไม่

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้มีความมุ่งหมายหลัก

1. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง
2. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง
3. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว
4. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
6. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
7. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งที่หนึ่ง

8. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิกสิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สอง

9. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิกสิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สาม

ความสำคัญของการศึกษา

ผลการศึกษาค้นคว้าทำให้ทราบว่า การตรวจให้คะแนนโดยกฎซิกสิงที่แตกต่างกัน 3 แบบ แบบใด ตำแหน่งใด ส่งผลต่อค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์สูงสุด อันจะเป็นแนวทางในการตรวจให้คะแนน และ คะแนนที่ได้จากการวัดที่น่าเชื่อถือที่สุด

ขอบเขตของการศึกษาค้นคว้า

ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย

ประชากรสำหรับสุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ จำนวน 444 โรงเรียน มีห้องเรียน 490 ห้องเรียน และมีนักเรียนจำนวน 9,684 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ จำนวน 35 ห้องเรียน มีนักเรียน จำนวน 790 คน ได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) และมีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit)

ตัวแปรที่ใช้ในการศึกษาค้นคว้า

1. ตัวแปรอิสระ

1.1 วิธีการตรวจให้คะแนนโดยกฎซิกสิงที่แตกต่างกัน 3 แบบ

1.1.1 วิธีการตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่อง

1.1.2 วิธีการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง

1.1.3 วิธีการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

1.2 ตำแหน่งการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด

2. ตัวแปรตาม คือ ความเชื่อมั่นของแบบทดสอบ

เนื้อหาวิชาที่ใช้ในการสร้างแบบทดสอบ

เป็นเนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 บทที่ 1 - 8 ภาคเรียนที่ 1 ปีการศึกษา 2543 ตามหลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533)

ข้อตกลงเบื้องต้น

1. ในการศึกษาครั้งนี้ นักเรียนต้องทำข้อสอบในแบบทดสอบทุกข้อ และจะนำเฉพาะข้อมูลที่มีการตอบผิดถึงตำแหน่งที่ 3 ของทั้ง 3 แบบมาวิเคราะห์
2. ข้อสอบที่นักเรียนตอบผิดตั้งแต่ข้อแรกจะไม่นำมาวิเคราะห์
3. ข้อสอบที่นักเรียนทำถูก หรือ ผิดหมดทุกข้อจะไม่นำมาวิเคราะห์

นิยามศัพท์เฉพาะ

1. จุดชิลลิ่ง หรือ ตำแหน่งการตอบผิด (Ceiling Point)

หมายถึง ตำแหน่งของข้อสอบข้อที่นักเรียนตอบผิดเป็นตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

2. กฎชิลลิ่ง (Ceiling Rules)

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ข้อสอบข้อที่นักเรียนตอบถูกหลังข้อที่ตอบผิดมีค่าความยากมากกว่า ถือว่านักเรียนตอบข้อสอบข้อนั้นโดยมีแนวโน้มที่จะเดามากกว่าใช้ความรู้จริง การศึกษาครั้งนี้มี

2.1 กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1)

2.2 กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2)

2.3 กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3)

3. กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1)

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิดเป็นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ซึ่งมีการกำหนดคะแนนในแต่ละตำแหน่ง ดังนี้

3.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิด (0) เป็นข้อแรกข้อที่ตอบถูกหลังจากนั้นไม่นำมานับคะแนน

3.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สามข้อ (000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4. กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2)

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิดเป็นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ซึ่งมีการกำหนดคะแนนในแต่ละตำแหน่ง ดังนี้

4.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิด (0) เป็นข้อแรกข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดเป็นข้อที่สอง(0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดเป็นข้อที่สาม (0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

5. กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3)

หมายถึง กฎของการตรวจให้คะแนนที่กำหนดความสามารถที่แท้จริงของผู้สอบจากข้อสอบข้อที่ตอบผิดเป็นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่

กำหนดตำแหน่งการตอบผิด เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด ซึ่งมีการกำหนดคะแนนในแต่ละตำแหน่ง ดังนี้

5.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

5.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สอง ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

5.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบ โดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สาม ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

6. การไม่กำหนดตำแหน่งการตอบผิด หรือการตรวจให้คะแนนแบบปกติ

หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

7. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

หมายถึง แบบทดสอบที่ผู้วิจัยสร้างขึ้น ประกอบด้วยชุดของข้อคำถามแบบปรนัยเลือกตอบชนิด 4 ตัวเลือก วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เนื้อหาบทที่ 1 - 8 ตามพฤติกรรมการเรียนรู้ของวิลสัน (Wilson 1971 : 645 - 696) ซึ่งวัดความรู้ด้านการคิดคำนวณ ความเข้าใจการนำไปใช้ และการวิเคราะห์จำนวน 50 ข้อ

8. ความยากของข้อสอบ (Item Difficulty : b)

หมายถึง ความสามารถของผู้สอบที่จะทำข้อคำถามถูกต้องที่ความน่าจะเป็น .50 ค่าความยากมีค่าระหว่าง $-\alpha$ ถึง $+\alpha$ แต่ในทางปฏิบัติจะมีค่าระหว่าง -2 ถึง +2 โดยค่า -2 แสดงว่าเป็นข้อสอบที่ง่ายมากและค่า +2 แสดงว่าเป็นข้อสอบที่ยากมาก ซึ่งในการศึกษาวิจัยครั้งนี้ผู้วิจัยประมาณค่าความยากโดยวิธีของราซช์ (Rasch Model)

9. ความเชื่อมั่นของแบบทดสอบ

หมายถึง ความสอดคล้องของคะแนนที่นักเรียนตอบข้อสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ในการศึกษาค้างนี้ประมาณจากสูตรสัมประสิทธิ์ r_B (Coefficient r_B) (บุญเขต ภิญโญอนันตพงษ์. 2538 : 51)

10. หลักฐานแสดงความเที่ยงตรงของแบบทดสอบ (Validity)

หมายถึง หลักฐานที่แสดงว่าแบบทดสอบวัดในสิ่งที่ต้องการวัดได้อย่างถูกต้องตรงตามความมุ่งหมาย ในการศึกษาค้างนี้เป็นการแสดงหลักฐานความเที่ยงตรงเชิงเนื้อหา (Content

Validity) ซึ่งผู้วิจัยสร้างให้ครอบคลุมเนื้อหาตามตารางวิเคราะห์รายละเอียดและประมาณค่าจากวิธีของโรวินเนลลี และแฮมเบิลตัน (บุญเชิด ภิญโญอนันตพงษ์. 2526 : 69 - 70 ; อ้างอิงมาจาก Rovineli Hambleton, 1977) คือให้ผู้เชี่ยวชาญ จำนวน 5 ท่านเป็นผู้ตรวจสอบ

11. การเดา

หมายถึง คะแนนข้อที่นักเรียนตอบถูกหลังข้อที่นักเรียนตอบผิดในแต่ละตำแหน่งของการตอบผิดตำแหน่งต่าง ๆ

12. ผู้เชี่ยวชาญ

หมายถึง ผู้ที่มีคุณวุฒิตั้งแต่ปริญญาโททางการวัดผลการศึกษา และ ครูผู้สอนวิชาคณิตศาสตร์เป็นเวลาไม่น้อยกว่า 5 ปีจำนวน 5 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาค้นคว้าครั้งนี้จะกล่าวถึงเอกสารและงานวิจัยดังต่อไปนี้ คือ

1. วิธีการตรวจให้คะแนน
2. การเรียงข้อคำถามในแบบทดสอบ
3. การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้
4. ความเชื่อมั่นของแบบทดสอบ
5. ทฤษฎีการตอบข้อสอบ (Item Response Theory)
6. งานวิจัยที่เกี่ยวข้องกับวิธีการตรวจให้คะแนน
7. งานวิจัยที่เกี่ยวข้องกับการเรียงข้อคำถามในแบบทดสอบ
8. สมมติฐานในการศึกษาค้นคว้า

1. วิธีการตรวจให้คะแนน

ปัจจุบันมีผู้ให้แนวคิดเกี่ยวกับวิธีการตรวจให้คะแนนแบบทดสอบชนิดต่าง ๆ เพื่อให้ได้คะแนนที่เป็นความสามารถที่แท้จริง หรือใกล้เคียงความสามารถที่แท้จริงของผู้สอบให้มากที่สุดไว้หลายวิธี ดังนี้

วิธีการตรวจให้คะแนนของคัมบัส (Coombs. 1956 : 308 - 313)

คัมบัสเสนอให้มีวิธีตอบ โดยให้เลือกคำตอบที่ผิดแทนการเลือกคำตอบที่ถูกพร้อมให้ความเห็นว่าการเลือกคำตอบที่ผิดซึ่งมีอยู่หลายตัวเลือกในคำถามหนึ่ง ๆ นั้น ผู้ตอบต้องใช้ส่วนของความรู้ย่อมมาพิจารณามากกว่าการเลือกคำตอบถูกเพียงคำตอบเดียว ความรู้ที่วัดได้จะสมบูรณ์มากกว่าเพราะผู้ตอบรู้ว่าสิ่งใดผิด สิ่งใดถูกวิธีนี้สามารถวัดสัดส่วนของความรู้ได้ละเอียดมากกว่าวิธีธรรมดา สำหรับการตรวจให้คะแนนนั้นคัมบัสให้ผู้ตอบทำเครื่องหมาย (X) บนตัวเลือกที่เห็นว่าผิด ถ้าไม่รู้ว่าถูกหรือผิดให้ข้ามไปไม่ต้องทำเครื่องหมายใด ๆ และจะให้ 1 คะแนนสำหรับผู้ตอบที่ทำเครื่องหมาย (X) ลงบนตัวเลือกที่ผิดจริง และให้ $1 - K$ (K คือจำนวนตัวเลือก) สำหรับผู้ตอบที่ทำเครื่องหมาย (X) ลงบนตัวเลือกที่ถูก คือถ้ามี 4 ตัวเลือกคะแนนที่ได้ในแต่ละข้อจะอยู่ระหว่าง - 3 ถึง + 3 คะแนน การให้คะแนนแบบนี้ยังมีข้อบกพร่องคือ ผู้ตอบที่มีความรู้ไม่เท่ากันมีโอกาสได้คะแนนเท่ากัน เช่นในแบบทดสอบที่มี 5 ตัวเลือก (มีคำตอบผิด 4 ตัวเลือก) ผู้ที่รู้ว่าตัวเลือกใดผิด ตัวเลือกใดถูก จะสามารถทำเครื่องหมาย (X) ลงบนตัวเลือกที่ผิดได้อย่างถูกต้องเมื่อรวมคะแนนของข้อนี้จะได้ 4 คะแนน ส่วนผู้ตอบอีกคนหนึ่งรู้เฉพาะตัวเลือกที่ถูกก็จะ

ทำเครื่องหมาย (X) ลงบนตัวเลือกอื่น ๆ ที่ไม่ใช่ตัวถูก ทั้งๆ ที่ไม่รู้ว่าตัวเลือกเหล่านั้นผิดหรือไม่ ผลการตรวจก็จะได้ 4 คะแนนเท่ากันโดยที่ทั้งสองคนนี้มีความรู้ไม่เท่ากัน

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

เดรสเซล และชมิทเสนอวิธีให้คะแนนความมั่นใจ วิธีนี้ผู้ตอบจะต้องบอกถึงระดับความมั่นใจของตัวเลือกที่เลือกว่าถูกต้องเพียงไร วิธีบอกความมั่นใจของเดรสเซล และชมิทมีการแบ่งระดับความมั่นใจละเอียดเป็น 5 ระดับโดยใช้กับแบบทดสอบเลือกตอบ 4 แบบ คือ 1) แบบเลือกตอบเสรี (Free Choice Test) เป็นแบบเลือกตอบที่ผู้สอบจะเลือกคำตอบที่ข้อก็ได้ ตามที่ตนคิดว่าถูกต้อง 2) แบบสอบถามความมั่นใจ (Degree of Certainly Test) เป็นแบบทดสอบที่ผู้สอบเลือกตอบได้หนึ่งข้อและบอกว่าตนเองแน่ใจในคำตอบนั้นอยู่ในระดับใดโดยแบ่งความมั่นใจออกเป็น 4 ระดับ 3) แบบมีคำตอบถูกมากกว่าหนึ่ง (Multiple Answer Test) เป็นแบบทดสอบที่มีจำนวนคำตอบถูกในแต่ละข้อไม่จำกัด ผู้สอบจะต้องทำเครื่องหมายทุกข้อที่ตนคิดว่าถูก 4) แบบมีคำตอบถูกมากกว่า 2 คำตอบ (Two Answer Test) จาก 5 ตัวเลือกในแต่ละข้อ ดังตาราง 1, 2

ตาราง 1 วิธีการให้คะแนนแบบเลือกเสรี

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

ตาราง 2 วิธีให้คะแนนแบบบอกระดับความมั่นใจ

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

วิธีการตรวจให้คะแนนของเดวิส และ ฟิฟเฟอร์ (Davis and Fiffer. 1959 : 159 - 169)

เดวิส และ ฟิฟเฟอร์ เสนอให้มีการตรวจให้คะแนนแบบให้น้ำหนักแก่ทุกตัวเลือกไม่เท่ากัน โดยให้ผู้เชี่ยวชาญหลาย ๆ ท่าน เป็นผู้กำหนดน้ำหนัก ซึ่งเป็นวิธีหนึ่งที่จะช่วยให้คะแนนจากแบบทดสอบมีความเชื่อมั่นที่สูงขึ้น แต่ยุ่งยากในทางปฏิบัติจึงไม่เหมาะสมที่จะนำมาใช้ในรูปแบบทดสอบที่ครูสร้างขึ้น

วิธีตรวจให้คะแนนของเฟรดเดอริก บี เดวิส (Davis. 1966 : 79)

เฟรดเดอริก บี เดวิส เสนอวิธีแก้การเดาโดยใช้สูตร

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

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

สูตรดังกล่าวเป็นวิธีแก้เดาโดยการหักสัดส่วนของจำนวนข้อที่ตอบผิดออกจากจำนวนข้อที่ตอบถูก วิธีนี้เคยถูกวิจารณ์ว่ามีรากฐานมาจากข้อตกลงเบื้องต้นที่ยังเป็นปัญหา คือ คำตอบผิดทุกข้อถือว่าเป็นผลเนื่องมาจากการเดา จากข้อเท็จจริงที่ว่าข้อสอบที่ตอบผิดนั้น บางครั้งผู้สอบไม่ได้เดา แต่เป็นเพราะได้รับความรู้ที่ผิด หรือข้อสอบนั้นมีตัวลวงที่มีเหตุผลดี (เพ็ญศรี สว่างเนตร. 2520 : 2)

วิธีการตรวจให้คะแนนของอีเบล (Ebel. 1965 : 224)

อีเบล เสนอสูตรการให้คะแนนเพื่อแก้โอกาสที่จะตอบถูกโดยบังเอิญขึ้นใหม่ คือ แทนที่จะหักคะแนนข้อที่ตอบผิดกลับเพิ่มคะแนนเท่ากับข้อที่ไม่ตอบ โดยถือว่าข้อที่เว้นไว้นั้นเป็นข้อที่ผู้ตอบไม่รู้ หรือยังไม่ทันแต่ถ้าเขาตอบโดยการเดา หรือทำเครื่องหมายลงในกระดาษคำตอบ โดยไม่ได้อ่านข้อสอบแล้วมีโอกาที่จะตอบถูกหนึ่งในจำนวนตัวเลือกของข้อนั้น โดยใช้สูตรแก้การเดาเป็น

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

เมื่อ	X_c	แทน	คะแนนที่แก้การเดา
	R	แทน	จำนวนข้อที่ตอบถูก

- C แทน จำนวนตัวเลือกของข้อสอบแต่ละข้อ
O แทน จำนวนข้อที่เว้น

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 ดังตาราง 3

ตาราง 3 วิธีให้คะแนนแบบกิบบอนส์ โอล์คิน และโซเบล

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

จะเห็นว่าการตอบและวิธีการให้คะแนนตามวิธีของกิบบอนส์และคณะ สามารถวัดคนที่มีความรู้จริงๆ ไปจนถึงคนที่มีความรู้จริงในเรื่องนั้น ๆ ทั้งยังสามารถแก้การเดาสุ่มได้อีกด้วย

วิธีการตรวจให้คะแนนโดยกฎขิลลิ่ง (Ceiling Rules)

จากการศึกษางานวิจัยในต่างประเทศและในประเทศ มีผู้ศึกษาวิจัยเกี่ยวกับผลการใช้กฎขิลลิ่งเพื่อหาความสามารถที่แท้จริงของผู้สอบ ดังนี้

1. สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) ได้ทำการศึกษาวิจัยเปรียบเทียบผลการใช้กฎขิลลิ่งจากการกำหนดตำแหน่งการตอบผิดต่างกัน ต่อค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์โดยใช้ความยากของข้อสอบเป็นเกณฑ์ในการ

พิจารณาการเดาตอบถูกของผู้สอบ เพราะผู้สอบแต่ละคนจะมีความสามารถระดับหนึ่งที่จะทำข้อสอบข้อที่มีความยากระดับต่าง ๆ ได้ถูกต้องตามขีดความสามารถของแต่ละคน ดังนั้นเมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุดตามรูปแบบของกัณฑ์แล้ว นักเรียนที่ตอบข้อสอบข้อที่ยากได้ถูกต้องในขณะที่ตอบข้อสอบข้อที่ง่ายผิด ความน่าจะเป็นในการตอบถูกนั้นควรเป็นไปโดยการเดามากกว่าตอบถูกด้วยความสามารถที่แท้จริง จึงไม่ควรได้คะแนนในข้อที่ตอบถูกนั้น แต่การตอบถูกในข้อต่อไปซึ่งเป็นข้อที่ยากกว่า จะสรุปโดยทันทีไม่ได้ว่านักเรียนทุกคนตอบถูกโดยการเดา เพราะนักเรียนบางคนก็ตอบข้อสอบข้อที่ง่ายผิดแต่ตอบข้อสอบข้อที่ยากกว่าถูกนั้นอาจมีสาเหตุมาจากความบกพร่องในการคิด เช่นบางคนคิดมากในข้อสอบบางข้อหรืออาจเกิดจากความสะเพร่าทำให้ตอบผิดทั้ง ๆ ที่มีความรู้ จึงทำให้การประมาณค่าความสามารถที่แท้จริงไม่ถูกต้องนัก และเป็นการไม่ยุติธรรมสำหรับนักเรียน สมบูรณ์ สุริยวงศ์ จึงเสนอวิธีการตรวจให้คะแนนโดยกฎชิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือ การตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้า กับการไม่กำหนดตำแหน่งการตอบผิดมากำหนดความสามารถที่แท้จริงของผู้สอบมีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

1.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิด (0) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สามข้อ (000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.4 การตอบผิดตำแหน่งที่สี่ หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สี่ข้อ (0000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.5 การตอบผิดตำแหน่งที่ห้า หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่ห้าข้อ (00000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.6 การไม่กำหนดตำแหน่งการตอบผิดหมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อสุดท้าย

2. วิลสัน และคณะ (Wilson and others. 1988 : 213 - 216) ศึกษาผลการใช้กฎซิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน ต่อค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์ โดยใช้ระดับความยากของข้อสอบเป็นเกณฑ์ในการพิจารณาการเดาตอบถูกของผู้สอบ เพราะผู้สอบแต่ละคนจะมีความสามารถระดับหนึ่งที่จะทำข้อสอบข้อที่มีความยากระดับต่าง ๆ ได้ถูกต้องตามขีดความสามารถของแต่ละคน ดังนั้นเมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุดตามรูปแบบของกัตแมนนักเรียนที่ตอบข้อสอบข้อที่ยากได้ถูกต้องในขณะที่ตอบข้อสอบข้อที่ง่ายผิดความน่าจะเป็นในการตอบถูกนั้นควรเป็นไปโดยการเดามากกว่าตอบถูกด้วยความสามารถที่แท้จริง นักเรียนจึงไม่ควรได้คะแนนในข้อที่ตอบถูกนั้นแต่การตอบถูกในข้อต่อไปซึ่งเป็นข้อที่ยากกว่าจะสรุปโดยทันทีไม่ได้ว่านักเรียนทุกคนตอบถูกโดยการเดา เพราะนักเรียนบางคนที่ตอบข้อสอบข้อที่ง่ายผิดแต่ตอบข้อสอบข้อที่ยากกว่าถูก อาจมีสาเหตุมาจากความบกพร่องในการคิด เช่น นักเรียนบางคนคิดมากในข้อสอบบางข้อหรืออาจเกิดจากความสะเพร่าทำให้ตอบผิดทั้ง ๆ ที่มีความรู้ จึงทำให้การประมาณค่าความสามารถที่แท้จริงไม่ถูกต้องนักและเป็นการไม่ยุติธรรมสำหรับนักเรียน วิลสัน และคณะจึงเสนอการตรวจให้คะแนนโดยใช้กฎซิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือ การตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้า กับการไม่กำหนดตำแหน่งการตอบผิดมากำหนดความสามารถที่แท้จริงของผู้สอบ มีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

2.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิด (0) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สามข้อ (000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.4 การตอบผิดตำแหน่งที่สี่ หมายถึง การกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่สี่ข้อ (0000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.5 การตอบผิดตำแหน่งที่ห้า หมายถึง การกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดติดต่อกันตั้งแต่ห้าข้อ (00000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.6 การไม่กำหนดตำแหน่งการตอบผิดหมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อสุดท้าย

3. จินตนา ต่อทรัพย์สินชัย (2533) อัมพร ยิ้มยิ้ม (2533) และมานะ คิวเจริญ (2539) ศึกษาวิธีการตรวจให้คะแนนโดยกฎซิลลิง ของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 613 - A) วิลสัน และคณะ (Wilson and others 1988 : 213 - 216) แล้วจึงทำการศึกษาวิจัยกับนักเรียนชั้นมัธยมศึกษา เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่เรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด โดยกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้ากับการไม่กำหนดตำแหน่งการตอบผิด มากำหนดความสามารถที่แท้จริงของผู้สอบ มีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

3.1 การตอบผิดตำแหน่งที่หนึ่ง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิด (0) เป็นข้อแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.2 การตอบผิดตำแหน่งที่สอง หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สอง (0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.3 การตอบผิดตำแหน่งที่สาม หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สาม (0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.4 การตอบผิดตำแหน่งที่สี่ หมายถึงการกำหนดให้คะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่สี่ (0,0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.5 การตอบผิดตำแหน่งที่ห้า หมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อที่ตอบผิดเป็นข้อที่ห้า (0,0,0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

3.6 การไม่กำหนดตำแหน่งการตอบผิดหมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 ถึงจนถึงข้อสุดท้าย

4. จากการศึกษาเกี่ยวกับการตรวจให้คะแนนโดยกฎซิลลิงในประเทศและต่างประเทศ ทำให้ผู้วิจัยเกิดความคิดว่าวิธีการตรวจให้คะแนนที่ใช้อยู่ในปัจจุบันซึ่งเป็นวิธีการให้คะแนนแบบ 0 - 1 ไม่มีการกำหนดตำแหน่งที่การตอบผิดน่าที่จะมีความคลาดเคลื่อนมาก เพราะเป็นการปฏิเสธคะแนนที่เป็นจริง (Type I error) และยอมรับคะแนนที่เป็นเท็จ (Type II error) เข้ามาเป็น

คะแนนรวม อย่างไรก็ตาม กฎชิลลิงของสมบูรณ์ สุริยวงศ์ วิลสันและคณะค่อนข้างจะเป็นวิธีการเข้มงวดมากจนเกินไป ผู้วิจัยเห็นว่าน่าจะเปิดโอกาสการตอบผิดที่ต่อเนื่องกันหลาย ๆ ครั้งเป็นจุดตัดเพื่อหาความสามารถที่เป็นคะแนนจริงได้มากขึ้น จึงคิดที่จะศึกษาค่าความเชื่อมั่นของแบบทดสอบที่เปิดโอกาสการตอบผิดที่ต่อเนื่องกันได้มากกว่า 1 ครั้งติดต่อกัน จากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง คือการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม ตำแหน่งที่สี่ ตำแหน่งที่ห้ากับการไม่กำหนดตำแหน่งการตอบผิด มากำหนดความสามารถที่แท้จริงของผู้สอบเมื่อเรียงข้อสอบจากข้อง่ายที่สุดไปหาข้อที่ยากที่สุด มีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

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

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

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

4.4 การตอบผิดตำแหน่งที่สี่ หมายถึง การกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่สี่ ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.5 การตอบผิดตำแหน่งที่ห้า หมายถึง การกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อที่ตอบผิดติดต่อกันมากกว่า 1 ข้อ (00) เป็นครั้งที่ห้า ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

4.6 การไม่กำหนดตำแหน่งการตอบผิดหมายถึงการกำหนดคะแนนรวมของแบบทดสอบโดยตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อ 1 จนถึงข้อสุดท้าย

วิธีการตรวจให้คะแนนของเกริกชัย ฮวบเจริญ (2525 : 5 - 6)

วิธีนี้ผู้สอบสามารถรู้ได้ทันทีว่าคำตอบที่ตนตอบ ถูก หรือ ผิด เพราะตอบโดยใช้เครื่องหมายกากบาทและปากกาถูกลิ้นจิ้มลงไปในช่องวงกลมที่มีอักษร ก ข ค ง และ จ คือถ้าผู้สอบเลือกคำตอบที่ถูกต้องปากกาก็จะสามารถจิ้มลึกลงไปถึงแผ่นล่างได้ ถ้าจิ้มไม่ถึงแผ่นล่างแสดงว่าเป็นคำตอบที่ไม่ถูกต้อง ผู้สอบจะทำข้อเดิมต่ออีกหรือไม่ก็ได้ ถ้าจะทำข้อเดิมต่อก็ให้กลับไปอ่าน

คำถามแล้วเลือกจิ้มในช่องวงกลมที่เหลือถ้ายังไม่ถูกอีกอาจตัดสินใจเลือกหรือไม่เลือกอีกก็ได้
ทำเช่นนี้เรื่อยไป การตรวจให้คะแนนใช้วิธี ดังตาราง 4

ตาราง 4 วิธีการตรวจให้คะแนนของเกริกชัย ฮวบเจริญ

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

วิธีการตรวจให้คะแนนของสำราญ มีแจ้ง (2525 : 10 - 11)

สำราญ มีแจ้ง เสนอวิธีตอบและตรวจให้คะแนน โดยให้ผู้ตอบเลือกตอบตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียวสำหรับข้อที่มั่นใจว่าถูกแน่ ส่วนข้อที่ผู้ตอบไม่มั่นใจก็ให้ตอบได้ 2 ตัวเลือก หรือ 3 ตัวเลือกที่คิดว่าถูก สำหรับการตรวจให้คะแนนจะให้คะแนนลดลงตามจำนวนตัวเลือกที่ตอบ ถ้าตอบเพียงตัวเลือกเดียวแล้วถูกจะได้ 4 คะแนน ถ้าผิดจะได้ 0 คะแนน ถ้าตอบ 2 หรือ 3 ตัวเลือกแล้วถูก (มีตัวเลือกถูกอยู่ในชุดที่เลือก) ก็จะได้คะแนน 2 และ 1 คะแนนตามลำดับ ถ้าผิดจะได้ 0 คะแนน ดังตาราง 5

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 ผู้ตอบที่มีความรู้น้อยจะเริ่มได้คะแนนก็ต่อเมื่อเขาสามารถกำจัดตัวลงได้อย่างน้อย 2 ตัว และเมื่อกำจัดตัวลงได้มากขึ้นคะแนนที่ได้จะเพิ่มเป็น 2 เท่า ของคะแนนเดิม วิธีการให้คะแนนของสำราญ มีแจ้งจึงไม่สามารถวัดคนที่มีความรู้น้อยได้จริงเพราะคนที่สามารถกำจัดตัวลงได้ 1 ตัวก็แสดงว่าเขามีความรู้ส่วนหนึ่งแล้วนอกจากนี้ยังทำให้ผู้ที่ยังคลุมเครือระหว่าง 3 และ 2 ตัวเลือกว่าตัวเลือกใดถูกเสี่ยงตอบเพียงตัวเลือกเดียว เพราะถ้าตอบถูกเขาจะได้คะแนนเพิ่มขึ้นจากการเสี่ยงเป็น 4 เท่า และ 2 เท่าของคะแนนที่เขาควรจะได้ตามลำดับทั้งยังเปิดโอกาสให้ผู้ตอบที่เดาสุ่มได้คะแนนอีกด้วย

วิธีการตรวจให้คะแนนของอนันต์ ศรีโสภา (2526 : 13 - 19)

อนันต์ ศรีโสภา เสนอวิธีการตอบวิธีใหม่โดยให้ตอบทุกตัวเลือกการตอบวิธีนี้ให้ผู้สอบระบุทั้งคำตอบที่ถูก และคำตอบที่ผิด เช่น ให้ทำเครื่องหมาย + ลงบนตัวเลือกที่เห็นว่าถูก และทำเครื่องหมาย - ลงบนตัวเลือกที่เห็นว่าผิด ถ้าตัวเลือกใดที่ไม่รู้ว่าถูก หรือผิดให้ข้ามไปไม่ต้องทำเครื่องหมายใด ๆ การตรวจให้คะแนน ถ้าผู้สอบทำเครื่องหมาย + ลงบนตัวเลือกที่ถูก หรือเครื่องหมาย - ลงบนตัวเลือกที่ผิดจะได้ตัวเลือกละ 1 คะแนน แต่ถ้าทำเครื่องหมาย + ลงบนตัวเลือกที่ผิด หรือเครื่องหมาย - ลงบนตัวเลือกที่ถูก จะได้ตัวเลือกละ -1 คะแนน ส่วนตัวเลือกที่ผู้สอบไม่ได้ทำเครื่องหมายใด ๆ เลยจะได้ 0 คะแนน ซึ่งการตรวจให้คะแนนวิธีนี้เป็นการวัดความรู้ในข้อคำถามนั้น ๆ ได้สมบูรณ์ยิ่งกว่าการเลือกคำตอบที่ถูกเพียงคำตอบเดียว และยังสามารถทราบความรู้ที่ผิดของผู้สอบได้ดีกว่าวิธีตอบแบบธรรมดาอีกด้วย

วิธีการตรวจให้คะแนนแบบธรรมดา (The Conventional Method)

บุญชม ศรีสะอาด (2540 : 122) กล่าวถึงการตรวจให้คะแนนตามวิธีธรรมดาว่าการให้ 1 คะแนนสำหรับคำตอบถูก และ 0 คะแนนสำหรับคำตอบผิด วิธีนี้จึงเรียกอีกอย่างหนึ่งว่า “วิธี 0 - 1” หรือ “วิธีให้ข้อถูกต้อง 1 คะแนน” หรือ “วิธีให้คะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิม” แต่เนื่องจากเป็นวิธีที่ใช้กันโดยทั่วไปในโรงเรียนและสถาบันการศึกษาจนถือเป็นปกติธรรมดาจึงเรียกว่าเป็น “วิธีธรรมดา” หรือ “วิธีประเพณีนิยม” การตรวจให้คะแนนวิธีนี้จะให้ 1 คะแนน ถ้าเป็นคำตอบที่ถูกต้องไม่ว่าข้อนั้นจะมีค่าความยากง่าย หรือมีความสำคัญมากน้อยเพียงใด ผู้ตอบมีโอกาสที่จะเดาแต่ละข้อถูกไม่เท่ากันขึ้นอยู่กับทราบคำตอบที่ถูกและผิดในข้อนั้น ๆ

วิธีให้คะแนนตามน้ำหนักที่กำหนดโดยผู้เชี่ยวชาญ

บุญชม ศรีสะอาด (2540 : 122) กล่าวว่าวิธีนี้จะให้ผู้เชี่ยวชาญหลาย ๆ ท่านกำหนดว่าแต่ละข้อควรตรวจให้คะแนนกี่คะแนน แล้วนำเอาผลการพิจารณาของผู้เชี่ยวชาญเหล่านั้นมาหาค่าเฉลี่ยเป็นคะแนนประจำข้อของข้อนั้น การพิจารณาของผู้เชี่ยวชาญจะถือเอาความสำคัญมากน้อยของข้อสอบเป็นเกณฑ์กำหนดคะแนน นอกจากวิธีนี้แล้ว ยังมีวิธีให้คะแนนที่ละเอียดโดย

ผู้เชี่ยวชาญกำหนดน้ำหนักคะแนนเป็นรายตัวเลือก ใครตอบตัวเลือกใดจะได้คะแนนประจำตัวเลือกนั้น

วิธีให้คะแนนตามความยากของข้อสอบ (Item Difficulty)

บุญชม ศรีสะอาด (2540 : 122) กล่าวถึงการถือหลักที่ว่าคนทำข้อยากได้ถูกจะมีความสามารถสูงกว่าคนทำข้อที่ง่ายได้ถูก การทำข้อยากได้ถูกจึงต้องได้คะแนนมากกว่าทำข้อที่ง่ายได้ถูก ดังนั้นการตรวจให้คะแนนตามความยากของข้อสอบแต่ละข้ออาจทำได้หลายวิธี เช่น

1. วิธีให้คะแนนตามค่าความยากมาตรฐาน ค่าความยากมาตรฐานเรียกว่า ค่าเดลต้า เขียนแทนด้วยสัญลักษณ์ Δ มีค่าตั้งแต่ 1 - 25 ถ้าข้อสอบยากจะมีค่าเดลต้าสูง ข้อสอบง่ายจะมีค่าเดลต้าต่ำ การหาค่าเดลต้าอาจหาโดยนำเอาค่าสัดส่วนของคนตอบถูกในกลุ่มสูงและสัดส่วนของคนตอบถูกในกลุ่มต่ำ ไปเปิดตาราง ของ จุง เตห์ ฟาน (Chung - Teh Fan) อ่านค่าเดลต้าออกมา หรือหาจากสูตรของ Henrysson ดังนี้

$$\Delta_i = 13 + 4Z_i$$

เมื่อ Z_i แทน ความเบี่ยงเบนมาตรฐาน ณ จุดที่บอกอัตราส่วนของผู้ที่ตอบถูก และ ผิดของข้อที่ i

2. วิธีให้คะแนนตามค่า q เมื่อพูดถึงค่าความยากของข้อสอบ โดยทั่วไปจะหมายถึงค่า p ซึ่งก็คือสัดส่วนของผู้ที่ตอบข้อนั้นถูก เช่น ถ้าค่า $p = .80$ แสดงว่าถ้ามีผู้สอบทั้งหมด 100 คน จะมีผู้ตอบถูก 80 คน ข้อสอบที่ยากจะมีค่า p ต่ำ ข้อสอบที่ง่ายจะมีค่า p สูง (ซึ่งมีลักษณะที่ตรงกันข้ามกับกรณีของเดลต้า) เมื่อจะให้น้ำหนักคะแนนตามความยาก กล่าวคือ ข้อที่ยากจะมีคะแนนมากกว่าข้อที่ง่ายกว่า ก็ไม่อาจใช้ค่า p ได้ต้องใช้ค่า q แทน ซึ่งค่า $q = 1 - p$ เมื่อทราบค่า p ก็หา q ได้ การหาค่า p หาจากการเปิดตารางของ จุง เตห์ ฟาน หรือใช้จากสูตรคำนวณที่ว่าค่าความยาก = จำนวนคนตอบถูกทั้งหมดหารด้วยจำนวนคนสอบทั้งหมด การให้คะแนนตามวิธีนี้ใครตอบข้อใดถูกก็จะได้คะแนนเท่ากับ q ของข้อนั้น

3. วิธีให้คะแนนตามค่า β_i ค่า β เป็นค่าความยากของข้อสอบที่ได้จากการวิเคราะห์แบบทดสอบรายข้อตามรูปแบบของราสค์ (Rasch Model) ค่า β คำนวณจากสูตรของ Wright และ Stone ดังนี้

$$\beta_i = \bar{X} \sqrt{(1 + \sigma^2 / 1.7^2) \ln [(1 - P_i) / P_i]}$$

เมื่อ $\bar{X}$ แทน ค่าเฉลี่ยของคะแนน

σ^2 แทน ความแปรปรวนของคะแนน

P_i แทน ค่าความยากข้อที่ i

โดยปกติ β มีค่าระหว่าง - 2.5 ถึง +2.5 เพื่อจัดค่าติดลบสู่พจน์ สุกมลสันต์ใช้ค่า β_i ซึ่งเป็นค่าความยากมาตรฐานโดยแปลงจากค่า β ตามวิธีของค่า Δ ดังนี้

$$\beta_i = 13 + 4 \beta_i$$

โดยทั่วไป ค่า β_i จะมีค่าระหว่าง 3 ถึง 23 การให้คะแนนตามวิธีนี้ใครตอบข้อใดถูก ได้คะแนนเท่ากับ β_i ของข้อนั้น

จากวิธีการตรวจ การกำหนดคะแนนวิธีการต่าง ๆ ดังได้กล่าวมาข้างต้นพอสรุปได้ว่า มีวิธีให้คะแนนแบบธรรมดา (0 - 1) วิธีให้คะแนนตามความสามารถในการรู้คำตอบที่ผิดและถูก วิธีให้โอกาสตอบได้มากกว่า 1 ครั้งโดยให้คะแนนน้อยลง วิธีให้น้ำหนักความมั่นใจในการเลือกตอบ วิธีให้โอกาสตอบมากกว่า 1 ตัวเลือกโดยให้คะแนนน้อยลง และวิธีการให้คะแนนโดยกฎซิลลิง ผู้วิจัยมีความสนใจที่จะนำวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด มาใช้ในการศึกษาครั้งนี้

2. การเรียงข้อคำถามในแบบทดสอบ

จากการศึกษาค้นคว้างานวิจัยที่เกี่ยวกับการวัดผลและประเมินผลของต่างประเทศและในประเทศพบว่ามีผู้ศึกษา และเสนอแนวคิดเกี่ยวกับการเรียงลำดับข้อคำถามไว้ ดังนี้

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

1. เรียงตามลำดับความยาก (Arrangement in Order of Difficulty) เป็นวิธีเรียงลำดับข้อคำถามตามความยากของข้อคำถามโดยการคำนวณทางสถิติหรือจากการตัดสิน

การเรียงข้อสอบแบบนี้มีข้อดี คือ ผู้สอบจะพบปัญหาง่ายในตอนแรก และไม่หมดกำลังใจเมื่อพบปัญหาที่ยากขึ้น

ส่วนข้อเสียที่พบได้ คือ เนื้อหาและพฤติกรรมที่ต้องการวัดไม่อยู่รวมเป็นพวกเดียวกันผู้สอบต้องเปลี่ยนความคิดอย่างรวดเร็ว อาจหมดกำลังใจที่จะทำแบบทดสอบ เพราะเมื่อเริ่มทำก็พบข้อยาก ข้อต่อไปของแบบทดสอบก็น่าที่จะยากขึ้น

2. เรียงลำดับความยากหมุนเวียน (Arrangement in Cyclic Order of Difficulty) วิธีนี้มีจุดมุ่งหมายเพื่อกระตุ้นให้ผู้สอบอ่านข้อสอบทุกข้อ เพราะถ้าผู้สอบอ่านต่อไปอีกระยะก็หนึ่งจะพบข้อง่ายอีก การจัดแบบนี้มีประโยชน์แต่ผู้สอบต้องเปลี่ยนความคิดในแต่ละเรื่องอย่างรวดเร็ว

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

4. เรียงตามจุดมุ่งหมายของการวัด (Arrangement According to the Goals Measured) เป็นการจัดเรียงข้อสอบเป็นกลุ่มตามพฤติกรรมการวัด

การเรียงข้อสอบแบบนี้มีข้อดี คือครูสามารถตรวจสอบว่าวัดตรงตามจุดมุ่งหมายที่ต้องการวัดหรือไม่

ส่วนข้อเสียที่พบได้ คือ ในทางปฏิบัติจะไม่สามารถรวมกลุ่มของปัญหาจากเนื้อหาที่คล้ายกันเข้าไว้ด้วยกัน

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

นูนนาลลี (Nunnally. 1959 : 159) มีความเห็นว่า ไม่ว่าจะเป็นแบบทดสอบที่ใช้ในการทำนาย หรือวัดผลสัมฤทธิ์ทางการเรียน อาจเรียงลำดับข้อคำถามจากง่ายไปหายาก หรืออาจเรียงลำดับความยากง่ายสลับกันไปและข้อคำถามที่เรียงตามลำดับความยากนั้นจะต้องเป็นแบบทดสอบที่ทดลองสอบและหาความเชื่อมั่นมาแล้ว แต่ถ้ายังไม่ได้ทดสอบและปรับปรุงควรเรียงลำดับข้อคำถามแบบสุ่ม ซึ่งสแตนเลย์และฮอปกินส์ (Stanley and Hopkins. 1972 : 191) ก็มีความเห็นในทำนองเดียวกันว่า การเรียงลำดับข้อคำถามควรเริ่มต้นจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดซึ่งเป็นไปตามหลักจิตวิทยาเพราะถ้านักเรียนพบข้อคำถามข้อที่ยาก ๆ ในตอนต้นของแบบทดสอบอาจทำให้เกิดความท้อใจโดยเฉพาะนักเรียนในกลุ่มความสามารถปานกลางและต่ำ

กรอนลันด์ (Gronlund. 1977 : 104 - 106) เสนอแนวคิดในการจัดเรียงข้อคำถามไว้ว่า ควรจัดเรียงข้อคำถามประเภทเดียวกันไว้ด้วยกัน เช่น คำถามแบบอัตนัย คำถามแบบถูกผิด คำถามแบบเลือกตอบ และอื่น ๆ และในแต่ละประเภทควรจัดเรียงตามจุดมุ่งหมาย เช่น ให้ความรู้ ความจำ ความเข้าใจ การนำไปใช้ ฯลฯ โดยเรียงจากข้อที่ง่ายไปหาข้อที่ยาก

ชวาล แพร์ตกุล (2518 : 134) กล่าวว่าข้อสอบในแต่ละฉบับควรจะเรียงคำถามเริ่มจากข้อง่ายไปหาข้อยากเสมอ ไม่ควรเรียงคำถามตามลำดับการสอนก่อนหลัง ไม่ควรเรียงตามลำดับเนื้อหาวิชาที่สอน หรือ เรียงตามความยากง่ายของบทเรียนเลยทีเดียว และถ้าเป็นข้อคำถามที่เริ่มใช้เป็นครั้งแรกต้องนึกคาดคะเนเอาเองว่าข้อใดยากง่ายกว่ากันแล้วเรียงลำดับตามนั้น

อนันต์ ศรีโสภ (2520 : 141) กล่าวแนะว่าการจัดเรียงข้อสอบที่ดี ครูควรจัดเรียงจากข้อง่ายไปหาข้อยาก เพราะยังไม่มีหลักฐานยืนยันว่าการจัดเรียงข้อสอบที่มีเนื้อหาเดียวกันเข้าไว้ด้วยกันจะช่วยให้ค่าความเที่ยงตรง ค่าความเชื่อมั่นสูง

อุทุมพร จามรมาน (2536 : 54) เสนอหลักการจัดเรียงข้อคำถามในแบบทดสอบ ดังนี้

1. จัดข้อคำถามรูปแบบเดียวกันไว้ด้วยกันพร้อมทั้งเขียนคำชี้แจงในการตอบให้ผู้ตอบเข้าใจอย่างชัดเจน

2. จัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อยากที่สุด

3. จัดเรียงข้อคำถามตามเนื้อหาของบทเรียน

4. จัดเรียงข้อคำถามตามจุดมุ่งหมายในการสอบ หรือพฤติกรรมที่จะต้องวัด

5. จัดเรียงข้อคำถามโดยการสุ่ม

ข้อดีของการจัดเรียงข้อคำถามในแต่ละรูปแบบของอุทุมพร จามรมาน

1. การจัดเรียงข้อคำถามตามรูปแบบข้อสอบจะทำให้ผู้สอบไม่สับสน ว่าควรตอบอย่างไรในข้อนั้น ๆ

2. การจัดเรียงข้อคำถามจากข้อง่ายไปหาข้อยาก จะทำให้ผู้สอบมีกำลังใจในการทำแบบทดสอบ

3. การจัดเรียงข้อคำถามตามเนื้อหาและจุดมุ่งหมาย หรือพฤติกรรมที่ต้องการวัดจะทำให้ผู้สอบมีความรู้ในเนื้อหาอะไรบ้าง บรรลุตามจุดมุ่งหมายใดบ้าง หรือมีพฤติกรรมการเรียนรู้ตามที่กำหนดไว้ในด้านใดบ้าง

จากการศึกษาทำให้ทราบว่า รูปแบบการจัดเรียงข้อคำถามของแบบทดสอบสามารถทำได้หลายวิธี และวิธีที่นักการศึกษา นักวัดผลได้เสนอแนวคิดไว้มี การเรียงตามค่าความยาก เรียงตามลำดับเนื้อหาวิชา เรียงตามลำดับจุดมุ่งหมายในการวัด (เรียงตามของพฤติกรรมการเรียนรู้) และเรียงแบบสุ่มซึ่งในแต่ละวิธี มีข้อดีข้อเสียแตกต่างกันขึ้นอยู่กับวัตถุประสงค์ของผู้ใช้ สำหรับการศึกษาในครั้งนี้ผู้วิจัยสนใจศึกษาการจัดเรียงข้อคำถามตามค่าความยากที่เรียงจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

3. การเขียนข้อสอบวัดพฤติกรรมการเรียนรู้

มีผู้ศึกษาเกี่ยวกับการเขียนข้อสอบวัดพฤติกรรมการเรียนรู้ไว้หลายลักษณะ ในที่นี้จะเสนอแนวคิดในการเขียนข้อสอบวัดพฤติกรรมการเรียนรู้ด้านพุทธิพิสัยที่สำคัญ ๆ ดังนี้

บลูม และคนอื่น ๆ (ลัวัน สายยศและ อังคณา สายยศ 2539 : 41 - 44 ; อ้างอิงมาจาก Bloom and others. 1956.) ได้แบ่งพฤติกรรมการเรียนรู้ออกเป็น 6 พฤติกรรมย่อย ๆ ดังนี้

1. ความรู้ (Knowledge) เป็นความสามารถในการระลึกในสิ่งที่ได้เรียนรู้มาแล้ว หรือก็คือความจำซึ่งแบ่งเป็นข้อย่อย ๆ ได้ ดังนี้

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

1.1.1 ความรู้เกี่ยวกับศัพท์และนิยาม (Knowledge of Terminology) ความรู้แบบนี้เป็นสัญลักษณ์ ศัพท์ นิยามที่ตกลงกันไว้เพื่อใช้ในเรื่องใดเรื่องหนึ่งเพื่อให้เป็นความหมายที่สะดวก

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 วิพากษ์วิจารณ์ รวมทั้งวิธีแสวงหาความรู้และลำดับขั้นของเวลา แบ่งเป็น 5 อย่าง คือ

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

1.2.2 ความรู้เกี่ยวกับแนวโน้มและลำดับขั้น (Knowledge of Trends and Sequences) เป็นความสามารถในการจำ เพื่อหาส่วนที่เกี่ยวข้องกับการแสดงออกทางแนวโน้มและลำดับขั้นตอนในเรื่องใดเรื่องหนึ่ง

1.2.3 ความรู้เกี่ยวกับการจำแนกประเภท (Knowledge of Classifications and Categories) เป็นความจำในเรื่องจัดประเภท กลุ่มชุดของความรู้ในเนื้อหาวิชาที่เรียนรู้มา

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

1.2.5 ความรู้เกี่ยวกับวิธีการ (Knowledge of Methodology) เป็นลักษณะการจำวิธีการในการค้นหาความรู้ จำเทคนิค และกระบวนการต่างๆ ที่เคยเรียนมาแล้ว

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

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

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

2. ความเข้าใจ (Comprehension) หมายถึง ความสามารถในการแปลความ ตีความ และ ขยายความจากสื่อความหมายต่าง ๆ ที่ได้พบเห็น แบ่งส่วนย่อยออก 3 อย่าง คือ

2.1 การแปลความ (Translation) เป็นความสามารถในการถ่ายเทความหมาย จากภาษาระดับหนึ่งมาเป็นอีกระดับหนึ่งให้เข้าใจกันง่ายขึ้น

2.2 การตีความ (Interpretation) เป็นความสามารถในการสรุป การแปลความ มองภาพส่วนรวมมาเป็นใจความสั้น ๆ อย่างได้ความ

2.3 การขยายความ (Extrapolation) เป็นความสามารถในการทำนาย หรือ คาดคะเนข้อเท็จจริงล่วงหน้าโดยอาศัยแนวโน้มที่มีมาแล้ว

3. การนำไปใช้ (Application) เป็นความสามารถในการนำหลักวิชาไปใช้แก้ปัญหาใน สถานการณ์ใหม่

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

4.1 วิเคราะห์ความสำคัญ (Analysis of Elements) เป็นการวิเคราะห์ว่าสิ่งที่อยู่ นั้นอะไรสำคัญ หรือจำเป็น หรือมีบทบาทที่สุด ตัวไหนเป็นเหตุ ตัวไหนเป็นผล

4.2 วิเคราะห์ความสัมพันธ์ (Analysis of Relationships) เป็นการหาความสัมพันธ์หรือความเกี่ยวข้องของส่วนย่อยในปรากฏการณ์ หรือเนื้อหานั้นเพื่อนำมาอุปมาอุปไมย

4.3 วิเคราะห์หลักการ (Analysis of Organizational Principles) เป็นความสามารถที่จะจับเค้าเงื่อนของเรื่องราวเห็นว่ายึดหลักการใด มีเทคนิค หรือยึดคติใด

5. การสังเคราะห์ (Synthesis) เป็นความสามารถในการผสมส่วนย่อยต่าง ๆ เข้าด้วยกัน เพื่อให้เป็นสิ่งใหม่อีกรูปแบบหนึ่งมีคุณลักษณะโครงสร้างหรือหน้าที่ใหม่แปลกแตกต่างไป จากของเดิมแบ่งย่อยออกเป็น 3 อย่าง คือ

5.1 สังเคราะห์ข้อความ (Production of Unique Communication) เป็นความสามารถในการสังเคราะห์ข้อความโดยสื่อหรือโดยการพูด การเขียน การวิพากษ์วิจารณ์ หรือหา ข้อยุติบางประการ

5.2 สังเคราะห์แผนงาน (Production of Plans and Proposed Set of Operations) เป็นความสามารถด้านการวางแผนหรือการจัดกระทำการออกแบบหรือการกระทำ ในลักษณะเดียวกันนี้

5.3 สังเคราะห์ความสัมพันธ์ (Derivation of a Set of Abstract Relation) เป็นความสามารถในการนำนามธรรมย่อย ๆ มาสัมพันธ์กันเกิดเป็นทฤษฎี สมมติฐาน สูตรหรือกฎ

6. การประเมินค่า (Evaluation) เป็นความสามารถในการพิจารณาตัดสินลงสรุปเกี่ยวกับคุณค่าของความคิดทุกชนิดเพื่อเปรียบเทียบเกณฑ์หรือมาตรฐานที่กำหนดให้ แบ่งย่อยออกเป็น 2 อย่าง คือ

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

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

วิลสัน (Wilson 1971 : 645 - 696) นำแนวคิดในการวัดพฤติกรรมการเรียนรู้ด้านพุทธิพิสัย (Cognitive Domain) ของบลูม และคณะ มาพัฒนาต่อเป็นลำดับขั้นพฤติกรรมวิชาคณิตศาสตร์ 4 ระดับ ซึ่งในแต่ละระดับประกอบด้วยพฤติกรรมย่อยหลายชั้น ดังนี้

1. วัดความรู้ด้านการคิดคำนวณ (Computation) เป็นการวัดความสามารถด้านการระลึกถึงสิ่งที่ได้เรียนมาแล้วทั้งในด้านข้อเท็จจริง ศัพท์ นิยาม ตลอดจนความสามารถในการดำเนินการคิดโจทย์ปัญหาต่าง ๆ ไม่ยุ่งยากซับซ้อน ไม่ต้องอาศัยการตัดสินใจ ทั้งนี้รวมถึงโจทย์ปัญหาที่เหมือนกับตัวอย่าง หรือแบบฝึกหัดที่เคยทำมาแล้ว การวัดพฤติกรรมขั้นนี้ยังแบ่งออกเป็น 3 ชั้นย่อย ๆ ได้แก่

1.1 วัดความรู้ความจำเกี่ยวกับข้อเท็จจริง (Specific Facts) เป็นการวัดด้านความสามารถที่จะระลึกถึงข้อเท็จจริงต่าง ๆ ที่ได้เรียนมา ตลอดจนพื้นฐานต่าง ๆ ทางคณิตศาสตร์ที่นักเรียนสร้างสมมาเป็นเวลานานอีกด้วย

ตัวอย่าง การเขียนเซตเขียนได้ก็แบบ

คำถามข้อนี้ถ้านักเรียนจำได้ว่าการเขียนเซตนั้นมี 2 แบบคือ “แบบแจกแจงสมาชิก” และ “แบบบอกเงื่อนไข” นักเรียนก็สามารถตอบโจทย์ปัญหาคำถามนี้ได้

1.2 วัดความรู้ความจำเกี่ยวกับศัพท์ และนิยาม (Terminology) เป็นการวัดความสามารถในการระลึกถึงศัพท์ นิยามต่าง ๆ ที่ได้เรียนมาแล้วได้ โดยไม่ต้องอาศัยการคิดคำนวณแต่อย่างใด และไม่ต้องการความรู้อื่นมาช่วย

ตัวอย่าง เซตที่ไม่มีสมาชิกเลยเรียกว่า ...

นักเรียนจะตอบคำถามข้อนี้ได้ นักเรียนจะต้องจำได้ว่า เซตที่ไม่มีสมาชิก คือ เซตว่าง

1.3 วัดความรู้ความจำเกี่ยวกับการใช้กระบวนการคิดคำนวณ (Algorithms) เป็นการวัดความสามารถในการใช้ข้อเท็จจริง ศัพท์ หรือ นิยาม และ กระบวนการ ที่ได้เรียนมาคิดคำนวณตามกระบวนการที่ได้เรียนมาแล้ว ในที่นี้หมายถึง การดำเนินการตามลำดับขั้นตอนที่ครูเคยสอนมา เช่น ลำดับขั้นตอนในการหารยาว ลำดับขั้นตอนการหา ห.ร.ม. หรือ ค.ร.น.

ตัวอย่าง ให้ $A = \{\alpha, \beta, \gamma, \theta\}$ และ $B = \{\alpha, \phi, \theta\}$ จงหา $A \cap B$

โจทย์ข้อนี้วัดความรู้ความจำเกี่ยวกับการใช้กระบวนการคิดคำนวณ นักเรียนต้องนำนิยามของอินเตอร์เซกชันมาใช้ โจทย์ข้อนี้เป็นโจทย์คล้ายคลึงกับตัวอย่าง และคำถามก็ไม่ได้สลับซับซ้อนเพราะถามแต่อินเตอร์เซกชันอย่างเดียวเท่านั้น

2. วัดความเข้าใจ (Comprehension) เป็นการวัดความสามารถในการนำความรู้ที่ได้เรียนแล้วมาสัมพันธ์กับโจทย์หรือปัญหาใหม่ ตลอดจนความสามารถในการตีความ แปลความ และขยายความได้ การวัดพฤติกรรมขั้นนี้แบ่งออกเป็น 6 ขั้นตอนย่อย ๆ คือ

2.1 วัดความเข้าใจเกี่ยวกับมโนคติ (Concept) เป็นการวัดความสามารถในการนำข้อเท็จจริงที่มีอยู่มาประมวลเข้าเป็นมโนคติ ซึ่งมีมโนคตินั้นมีความซับซ้อนกว่าข้อเท็จจริง ต้องอาศัยความรู้ต่าง ๆ มาผสมผสานกัน คำถามเกี่ยวกับมโนคตินี้ ครูจะต้องไม่เคยบอก หรือ สอนมาก่อน เพราะว่าถ้าเคยบอกมาก่อนแล้วจะกลายเป็นวัดความรู้ความจำเกี่ยวกับข้อเท็จจริง

ตัวอย่าง จำนวนเต็มต่างกับจำนวนธรรมชาติอย่างไร

คำถามข้อนี้ ผู้ตอบจะต้องมีความรู้เกี่ยวกับข้อเท็จจริงในเรื่องเซตของจำนวนธรรมชาติและมองเห็นความสัมพันธ์ระหว่างเซต 2 เซต คือสามารถเห็นความสัมพันธ์ว่าเซตของจำนวนธรรมชาติเป็นสับเซตของจำนวนเต็ม

2.2 วัดความเข้าใจเกี่ยวกับ หลักการ กฎ และการสรุปอ้างอิงเป็นกรณีทั่วไปทางคณิตศาสตร์ (Principles, Rules and Generalization) เป็นการวัดความสามารถในการนำเอาหลักการ กฎ และ ความรู้เกี่ยวกับมโนคติ ไปสัมพันธ์กับโจทย์ปัญหาจนได้แนวทางในการแก้ปัญหาโจทย์ได้

ตัวอย่าง ถ้าเลือกจุดทศนิยมของจำนวนหนึ่งไปทางขวา 3 ตำแหน่ง

หมายความว่าอย่างไร

- ก. คูณจำนวนนั้นด้วย 3
- ข. หารจำนวนนั้นด้วย 100
- ค. คูณจำนวนนั้นด้วย 1,000
- ง. หารจำนวนนั้นด้วย 1,000

2.3 วัดความเข้าใจในโครงสร้างทางคณิตศาสตร์ (Mathematical Structure) ใน
ขั้นนี้คำถามที่ใช้วัดพฤติกรรมมักจะเป็นคำถามเกี่ยวกับศัพท์ และนิยามในคณิตศาสตร์แผนใหม่

ตัวอย่าง จำนวนที่หายไป คือ จำนวนใด

$$24 + 76 = 76 + [...]$$

- ก. 24
- ข. 34
- ค. 52
- ง. 100

2.4 วัดความสามารถในการแปลงโจทย์ปัญหา จากรูปแบบหนึ่งไปยังอีกรูปแบบ
หนึ่ง (Transform Problem Elements from One Mode to Another) เป็นการวัดความสามารถ
การแปลงข้อความที่กำหนดให้เป็นข้อความใหม่ หรือภาษาใหม่อีกูปหนึ่งซึ่งมีความหมายคงเดิม
เป็นต้นว่าเปลี่ยนโจทย์ปัญหาให้อยู่ในรูปของสมการ ซึ่งการวัดในขั้นนี้ไม่รวมถึงวิธีการในการหา
คำตอบจากสมการนั้น

ตัวอย่าง จำนวนคนซึ่งเป็นสมาชิกของ $A \cap B$ มี 69 คน ถ้าจำนวนสมาชิก
ของเซต A มี 41 คน และจำนวนสมาชิกของเซต B มี 53 คน จงเขียนแผนภาพเวนอยเลอร์
แสดง $A \cup B$

2.5 วัดความสามารถในการดำเนินความคิดตามแนวของเหตุผลที่วางไว้ (Follow
a Line of Reasoning) วัดความสามารถในการอ่าน และเข้าใจกับข้อความทางคณิตศาสตร์ และ
สามารถบอกได้ว่าผลสรุปในแต่ละขั้นมาจากเหตุผลใด

ตัวอย่าง นาย ก ชื่อนิมิต ช้อน และล้อม มาชุดหนึ่ง ซึ่งผู้ขายรับรองว่าทำด้วย
สแตนเลส หลังจากใช้ไปแล้วช่วงระยะเวลาหนึ่งเขาพบว่าของใช้ชุดนี้เริ่มเป็นสนิม เขาจึงคิดว่า
ของที่ซื้อมาไม่ใช่สแตนเลส และจะนำไปคืน จงใช้ข้อความที่กำหนดตอบคำถามต่อไปนี้

- ก. ข้อความที่ต้องพิสูจน์คืออะไร
- ข. สมมติฐานคืออะไร
- ค. ผลสรุปที่ได้จากสมมติฐานคืออะไร
- ง. ความจริงที่ปรากฏคือขัดกับสมมติฐาน

คำถามข้อนี้ถ้านักเรียนอ่านเข้าใจข้อความที่ให้มาและสามารถดำเนินความคิดตามแนวของเหตุผลแต่ละข้อได้ เช่น ในข้อ ก “ข้อความที่ต้องพิสูจน์คืออะไร” นักเรียนอ่านเข้าใจย่อมตอบได้ว่าจะต้องพิสูจน์ว่าของที่ซื้อมาเป็นสแตนเลสหรือไม่

- ข้อ ก ของที่ซื้อมาเป็นสแตนเลสหรือไม่
- ข้อ ข สมมติฐานก็คือของที่ซื้อมาเป็นสแตนเลส
- ข้อ ค ผลสรุปที่ได้จากสมมติฐานก็คือ ถ้าเป็นสแตนเลสเมื่อใช้แล้วจะไม่เป็นสนิม
- ข้อ ง ความจริงที่ปรากฏขัดกับสมมติฐานคือเมื่อใช้ไปแล้วเป็นสนิม

2.6 วัดความสามารถในการอ่านและตีความโจทย์ปัญหา (Read and Interpreter Problem) การวัดความสามารถในการอ่านและตีความจากโจทย์ปัญหาที่กำหนดให้ เพื่อทราบว่ามีโจทย์ต้องการอะไร โจทย์กำหนดอะไรให้บ้าง สิ่งที่โจทย์กำหนดให้ยังขาดส่วนใดบ้าง รวมทั้งการแปลความหมายจากกราฟ หรือข้อมูลทางสถิติ ตลอดจนการแปลสมการหรือตัวเลขเป็นรูปภาพ

ตัวอย่าง นายแดงซื้อสินค้าอย่างหนึ่ง ซึ่งมีป้ายติดราคาไว้ 60 บาท แต่จากตารางลดราคาให้ 20% เขาจะได้ส่วนลดเท่าใด จากโจทย์ข้างต้นนี้ จงตอบคำถามต่อไปนี้

- ก. อัตราการลดราคาเป็นเท่าไร
- ข. ฐานที่ใช้ในการลดราคาเป็นเท่าใด
- ค. โจทย์ต้องการให้คำนวณหาอะไร

3. วัดการนำไปใช้ (Application) เป็นการวัดความสามารถในการแก้ปัญหาต่าง ๆ ที่คล้ายคลึงกับที่เคยเรียนมาแล้วนั้นคือนักเรียนจะต้องผสมผสานความรู้ความสามารถจากขั้นที่ 1 และ 2 มาใช้แก้โจทย์ปัญหาซึ่งมีหลายขั้นตอนในการจัดกระทำเพื่อให้คำตอบออกมา ดังนั้นจึงมีความจำเป็นจะต้องมีการเลือกตัดสินใจว่าจะทำขั้นตอนใดก่อน หลัง พฤติกรรมขั้นนี้ยังแบ่งออกเป็น 4 ขั้นย่อย คือ

3.1 วัดความสามารถในการแก้โจทย์ปัญหาที่คุ้นเคย (Solve Routine Problem) หรือ ปัญหาที่เคยเรียนมาแล้ว พฤติกรรมขั้นนี้เป็นความสามารถในการแก้ปัญหาที่คล้ายคลึง

แต่ไม่ใช่ข้อเดียวกันกับตัวอย่างหรือแบบฝึกหัดที่นักเรียนเคยทำมาแล้ว นักเรียนจะต้องอาศัยความรู้ความจำเกี่ยวกับการคิดคำนวณ และความเข้าใจ มาผสมผสานกันแก้ปัญหา เพื่อให้ได้ผลลัพธ์ออกมา

ตัวอย่าง กำหนดให้ $U = (1, 2, 3, \dots, 10)$
 $P = (2, 3, 4, 6)$
 $Q = (1, 5, 7, 9)$
 จงหาค่าของ $P = (P \cup Q)$

โจทย์ข้อนี้จะเห็นว่า ถ้านักเรียนเข้าใจวิธียูเนียน และผลต่างของเซตสองเซต และคอมพลีเมนต์ของเซตมาแล้ว ก็สามารถใช้วิธีการเดียวกันนั้นมาคำนวณหาคำตอบออกมาได้

3.2 วัดความสามารถในการเปรียบเทียบ (Make Comparisons) เป็นการวัดความสามารถในการหาความสัมพันธ์ โดยการเปรียบเทียบข้อมูลที่โจทย์ให้มา 2 ชุด ในการแก้ปัญหา อาจจะต้องใช้วิธีการคำนวณ ความเข้าใจ แล้วจึงมาเปรียบเทียบเพื่อตัดสินใจ

ตัวอย่าง คะแนนต่อไปนี้ เป็นผลสอบวิชาคณิตศาสตร์ของนักเรียน 20 คน

15, 12, 19, 12, 13, 16, 10, 14, 11, 10

17, 16, 17, 14, 11, 12, 15, 20, 9, 15,

มีนักเรียนกี่คนที่ได้คะแนนสูงกว่าคะแนนเฉลี่ยเลขคณิตของกลุ่ม

ก. 3 คน

ข. 9 คน

ค. 10 คน

ง. 11 คน

โจทย์ข้อนี้นักเรียนที่เคยเรียนเรื่องการหาค่าคะแนนเฉลี่ยมาแล้ว ยังตอบไม่ได้ทันทีเพราะต้องคำนวณหาค่าคะแนนเฉลี่ยแล้วนำไปเปรียบเทียบกับข้อมูลที่ละตัว จึงจะตอบคำถามได้

3.3 ความสามารถในการวิเคราะห์ข้อมูล (Analyze Data) เป็นความสามารถในการจำแนก และตัดสินใจว่าข้อมูลส่วนใดจำเป็นหรือไม่จำเป็นต่อการแก้ปัญหาโจทย์

ตัวอย่าง ลูกบาศก์ลูกหนึ่งมีด้านยาวด้านละ 12 เซนติเมตร ถ้าลูกบาศก์อีกลูกหนึ่งมีความยาวแต่ละด้านเป็น 2 เท่าของลูกบาศก์เดิม จงหาว่าลูกบาศก์ใหม่ มีปริมาตรเป็นกี่เท่าของลูกบาศก์เดิม

ก. 2 เท่า

ข. 4 เท่า

ค. 6 เท่า

ง. 8 เท่า

นักเรียนจะทำโจทย์ข้อนี้ได้ต้องทราบว่า ปริมาตรของลูกบาศก์หาได้อย่างไร และการที่ความยาวของแต่ละด้านเพิ่มขึ้นเป็นสองเท่าของความยาวเดิมนั้นจะมีผลทำให้ปริมาตรเพิ่มขึ้นอย่างไรเมื่อเปรียบเทียบกับปริมาตรเดิม

3.4 ความสามารถในการมองเห็นแบบลักษณะโครงสร้างที่เหมือนกัน และการสมมาตร (Recognize Patterns, Isomorphisms, and Symmetries) พฤติกรรมในขั้นนี้เกี่ยวกับการระลึกถึงข้อมูล แปลงปัญหาการจัดกระทำข้อมูล การมองเห็นความสัมพันธ์ระหว่างข้อมูลที่คุ้นเคยกับข้อมูลที่กำหนดให้ หรือจากปัญหาที่กำหนดขึ้น

ตัวอย่าง ตัวเลขหลักหน่วยของ 4^{10} คือ ข้อใด

ก. 0

ข. 2

ค. 4

ง. 6

จ. 8

4. วัดการวิเคราะห์ (Analysis) เป็นการวัดความสามารถที่จะตอบปัญหาจากโจทย์ ปัญหาที่มีลักษณะซับซ้อน พลิกแพลง ซึ่งนักเรียนไม่เคยลองฝึกทำมาก่อน แต่ทั้งนี้มิได้หมายความว่าโจทย์ปัญหานั้นจะอยู่นอกเหนือขอบข่ายเนื้อหาวิชาที่เคยเรียนมา ดังนั้น การแก้โจทย์ปัญหาที่วัดพฤติกรรมในขั้นนี้ จึงครอบคลุมความรู้ความสามารถในสามขั้นที่กล่าวมาแล้วรวมทั้งการมีความคิดริเริ่มสร้างสรรค์ เพื่อสามารถค้นพบวิธีการ หรือแนวทางในการแก้โจทย์ปัญหานั้นได้ การวัดพฤติกรรมขั้นวิเคราะห์แบ่งออกเป็นการวัดพฤติกรรมย่อย ๆ อีก 5 พฤติกรรม คือ

4.1 วัดความสามารถในการแก้โจทย์ปัญหาที่ไม่คุ้นเคยมาก่อน (Solve Nonroutine Problems) คำถามในขั้นนี้เป็นคำถามที่ต้องคิดซับซ้อน เป็นโจทย์ที่ไม่ได้อยู่ในแบบฝึกหัด หรือตัวอย่างไม่เคยเห็นมาก่อน นักเรียนจะแก้โจทย์ได้ต้องอาศัยความคิดสร้างสรรค์ของคณิตศาสตร์ ต้องเข้าใจโมโนมิติ หรือ นิยามตลอดจนทฤษฎีต่าง ๆ ที่ครูสอนมาแล้วเป็นอย่างดี และใช้ความรู้เหล่านั้นมาผสมผสานกันแก้โจทย์

ตัวอย่าง กำหนดให้ P เป็นจำนวนเฉพาะที่เป็นบวก และ $P = r^2 - t^2$

จงหา t ถ้า ก. $r = 7$, ข. $r = 157$, ค. $r = 58$

คำถามข้อนี้ นักเรียนจะต้องไม่เคยเห็นมาก่อนจะเห็นว่านักเรียนจะต้องอาศัยความรู้ในเรื่องจำนวนเฉพาะ และ จะต้องรู้ว่าเมื่อแทนค่า $r = 7$ ลงในสมการที่กำหนดจะได้ $P = 49 - t^2$ เมื่อแยกตัวประกอบจะได้ $P = (7 - t)(7 + t)$ นักเรียนจะต้องรู้จักวิเคราะห์ว่า t มีค่าเท่าใดได้บ้าง ซึ่งพบว่า $t < 7$ เพราะถ้า $t \geq 7$ แล้วจะทำให้ P มีค่าเป็นศูนย์ และเป็นลบ ซึ่งเป็นไปไม่ได้เพราะขัดกับสิ่งที่กำหนดให้นักเรียนจะต้องวิเคราะห์ต่อไปว่า t เป็นอะไรได้บ้างที่น้อยกว่า 7 แล้วทำให้ P เป็นจำนวนเฉพาะ ก็จะพบว่า $t = 6$ จะทำให้ $P = 13$ ซึ่งเป็นจำนวนเฉพาะ ส่วนค่า t อื่นเมื่อตรวจสอบแล้วไม่เป็นจำนวนเฉพาะ ข้อ ข และ ข้อ ค ก็ใช้วิธีการคิดทำนองเดียวกันกับข้อ ก

4.2 วัดความสามารถในการค้นหาความสัมพันธ์ (Discover Relationships)

ข้อสอบที่วัดพฤติกรรมขั้นนี้ นักเรียนจะต้องจัดส่วนต่าง ๆ ที่โจทย์กำหนดให้ใหม่ แล้วสร้างความสัมพันธ์ขึ้นเพื่อใช้ในการแก้ปัญหา พฤติกรรมในขั้นนี้ที่ต่างจากขั้นนำไปใช้ตรงที่นักเรียนต้องรวบรวมสิ่งที่โจทย์กำหนดให้มาเป็นความสัมพันธ์ขึ้นใหม่แทนที่จะนำความสัมพันธ์อันเดิมที่เคยพบมาแล้วมาใช้กับข้อมูลชุดใหม่เท่านั้น

ตัวอย่าง ความยาวของเส้นทแยงมุมของสี่เหลี่ยมจัตุรัสเป็น $(X + Y)$ จงหาพื้นที่ของสี่เหลี่ยมจัตุรัส

คำถามข้อนี้ ต้องการให้นักเรียนค้นหาความสัมพันธ์ระหว่างความยาวของด้าน และความยาวของเส้นทแยงมุมของสี่เหลี่ยมจัตุรัสก่อน หลังจากนั้นจึงหาคำตอบได้โดยง่าย

4.3 วัดความสามารถในการสร้างข้อพิสูจน์ (Construct Proofs) เป็นการวัดความสามารถในการสร้างข้อพิสูจน์โจทย์ปัญหาที่ไม่เคยพบมาก่อน นักเรียนจะต้องอาศัยนิยามและทฤษฎี ความรู้ต่าง ๆ มาช่วยในการแก้ปัญหา

ตัวอย่าง จงพิสูจน์ว่า $x^2 + y^2 = x + y$ เมื่อ x และ y เป็นจำนวนเต็มบวก

4.4 วัดความสามารถในการวิพากษ์วิจารณ์ (Criticize Proof) เป็นการวัดการใช้เหตุผลที่ควบคู่ไปกับความสามารถในการพิสูจน์ ซึ่งต้องการให้นักเรียนมองเห็นหรือเข้าใจการพิสูจน์นั้นว่าถูกต้องหรือไม่ มีตอนใดผิดพลาดบ้าง

ตัวอย่าง ให้ a และ b แทนจำนวนเต็ม ซึ่ง $a = b = 1$

- 1) $(a - b)a = (a - b)(b + 1)$ เอา $(a - b)$ คูณทั้งสองข้าง
- 2) $a^2 - ab = ab + a - b^2 - b$ กฎการกระจาย
- 3) $a^2 - ab - a = ab + a - b^2 - b - a$ เอา $(-a)$ บวกทั้งสองข้าง

$$4) a(a - b - 1) = b(a - b - 1) \text{ เอาตัวร่วมออกนอกวงเล็บ}$$

$$5) a = b(a - b - 1) \text{ หารทั้งสองข้าง}$$

4.5 วัดความสามารถในการสร้างสูตรและการทดสอบความถูกต้องของสูตรซึ่งใช้เป็นกรณีทั่วไปได้ (Formulate and Validate Generalization) พฤติกรรมขั้นนี้จะคล้ายกับข้อที่ 4.3 แต่อาจซับซ้อนมากกว่า คือนักเรียนต้องสามารถสร้างสูตรขึ้นมาใหม่โดยให้สัมพันธ์กับเรื่องที่เคยเรียนมาแล้ว และต้องสมเหตุสมผลสามารถใช้ได้ในทุกกรณี

ล้วน สายยศ (2517 : 36 - 37) กล่าวถึงแบบทดสอบมาตรฐานทางคณิตศาสตร์ว่านิยมสร้างเพื่อวัดพฤติกรรมการเรียนรู้ทางคณิตศาสตร์ ดังนี้

1. คณิตศาสตร์ทักษะ หมายถึง ความคล่องแคล่วในการบวก ลบ คูณ หาร ถอดราก หรือทำเลขซ้อนได้อย่างถูกต้องรวดเร็วในเวลาที่กำหนดให้ ซึ่งไม่เกี่ยวข้องกับอิทธิพลของความเข้าใจด้านภาษา การเขียนข้อสอบจึงเป็นเรื่องของความสัมพันธ์ระหว่างตัวเลข หรือตัวแปรโดยอาศัยเครื่องหมายทางคณิตศาสตร์ตลอดเวลา

2. คณิตศาสตร์เหตุผล การออกข้อสอบคณิตศาสตร์แบบนี้ จะถามเน้นในเรื่องวิธีการหลักการ การแปลความ ตีความ ขยายความ การไล่เรียงหาเหตุผล การเปรียบเทียบความคิดริเริ่มสร้างสรรค์ มโนภาพ ขั้นตอนการพิสูจน์ และการประมาณค่า เป็นต้น ข้อสอบแบบนี้ความเข้าใจด้านภาษามีอิทธิพลอยู่เพราะข้อคำถามจะต้องใช้อธิบายด้วยภาษาเป็นส่วนใหญ่

3. คณิตศาสตร์โจทย์ปัญหา จุดมุ่งหมายในการออกข้อสอบคณิตศาสตร์โจทย์ปัญหาก็เพื่อประสงค์ให้นักเรียนแก้ปัญหาเป็น คือ สามารถนำความรู้ทางคณิตศาสตร์ไปใช้ได้มุ่งตรวจสอบผลสัมฤทธิ์สุดท้ายในการคิด คือ จะดูผลลัพธ์ว่าเป็นเท่าไร โดยไม่ทราบวิธีการหรือเหตุผล

ชวาล แพรัตกุล (2517 : 46 - 47) ชี้ให้เห็นว่าโครงสร้างของแบบทดสอบคณิตศาสตร์ในปัจจุบันนิยมแยกชุดข้อคำถามออกเป็น 3 ฉบับ คือ แบบทดสอบสำหรับวัดทักษะ หรือความคล่องแคล่วแม่นยำในการคำนวณตัวเลข แบบทดสอบสำหรับวัดความสามารถในการแก้โจทย์ปัญหา และแบบทดสอบสำหรับวัดความคิดรวบยอด หรือที่เรียกว่า แบบทดสอบวัดเหตุผลเชิงคณิตศาสตร์ ซึ่งมีรายละเอียด ดังนี้

1. แบบทดสอบประเภททักษะ แบบทดสอบชนิดนี้ต้องการวัดความคล่องแคล่วแม่นยำในการบวก ลบ คูณ หารตัวเลข ซึ่งเป็นสมรรถภาพพื้นฐานในการเรียนคณิตศาสตร์ เทียบได้กับความสามารถในการสะกดคำของวิชาภาษาไทย ฉะนั้นโครงสร้างของแบบทดสอบชนิดนี้จึงไม่มีอิทธิพลของภาษาเข้ามาเกี่ยวข้องด้วย จัดว่าเป็นการวัดความคล่องแคล่วในการคำนวณตัวเลขโดยเฉพาะ

2. แบบทดสอบประเภทโจทย์ปัญหา แบบทดสอบชนิดนี้ประกอบด้วยคำถามประเภทโจทย์ต่าง ๆ ทำนองเดียวกับโจทย์คณิตศาสตร์ทั่ว ๆ ไป แต่ลีลาคำถามมุ่งที่จะวัดความสามารถ

ในการแปลความหมาย และการหาความสัมพันธ์ของตัวเลขจำนวน กับวิธีการแก้ปัญหาต่าง ๆ เป็นสำคัญ คำถามประเภทโจทย์ปัญหานี้ไม่ได้หมายถึงคำถามที่ให้นักเรียนนำตัวเลขไปแทนค่าในสูตรโดยตรงดังเช่นในเรื่องการหาดอกเบี้ยหรือให้หาพื้นที่ของวงกลมเลยเพราะคำถามชนิดนี้ไม่มีลักษณะเป็นปัญหา แต่เป็นการวัดความสามารถในการจำสูตร และ กฎเกณฑ์ต่าง ๆ แล้วเอาตัวเลขจากโจทย์ไปแทนค่าในสูตรเท่านั้นเอง

3. แบบทดสอบประเภทเหตุผล แบบทดสอบชนิดนี้ต้องการวัดสมรรถภาพด้านความคิดรวบยอดเชิงคณิตศาสตร์ ซึ่งเกี่ยวข้องกับหลักการ วิธีการ และมโนภาพในความคิดแบบนามธรรม นั่นคือ ต้องการวัดว่าเมื่อเด็กได้เรียนรู้กฎเกณฑ์ต่าง ๆ แล้วเขาจะสามารถขยายความหมายออกเป็นรูปของภาษาหรือสัญลักษณ์ให้กว้างขวางจากเดิมได้หรือไม่ แบบทดสอบประเภทนี้จะช่วยให้ครู อาจารย์ได้ทราบว่าหลังจากเรียนเนื้อหาต่าง ๆ แล้ว นักเรียนเกิดความคิดในเรื่องความสัมพันธ์ของปริมาณเหล่านั้น หรือไม่ หรือ เพียงแต่จำสูตร และ วิธีการตามตำราได้โดยไม่มีจินตนาการในกฎเกณฑ์เหล่านั้นเลย

สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ (2538 : 58) ให้ความหมายทักษะการคิดคำนวณว่า หมายถึงความสามารถในการคิดหาคำตอบจากประโยคสัญลักษณ์ หรือโจทย์คำนวณเพื่อให้ได้คำตอบที่ถูกต้องและรวดเร็ว

ทบวงมหาวิทยาลัย (2524 : 99) เสนอแนะลักษณะการแสดงออกของพฤติกรรมการเรียนรู้ทางคณิตศาสตร์ไว้ ดังนี้

1. ทักษะการคิดคำนวณเป็นความสามารถในการการบวก ลบ คูณ หาร การแก้สมการอย่างง่าย ๆ ที่ไม่ซับซ้อน การแทนค่าตัวแปร หรือสัญลักษณ์ต่าง ๆ ได้อย่างคล่องแคล่วแม่นยำถูกต้อง และรวดเร็วในเวลาที่กำหนดให้

2. การให้เหตุผลทางคณิตศาสตร์เป็นความสามารถในการตีความ ขยายความ สรุปผล เปรียบเทียบบอกขั้นตอนก่อน หลัง เช่น บอกว่าคู่อันดับใดสอดคล้องกับสมการที่กำหนดให้ได้

3. การแก้ปัญหาทางคณิตศาสตร์ เป็นความสามารถในการแก้โจทย์ปัญหาต่าง ๆ ทั้งที่เคยพบเห็น และไม่เคยพบเห็นมาก่อน โดยอาศัยการคำนวณ การให้เหตุผล และมโนคติต่าง ๆ ประมวลเข้าด้วยกันในการแก้ปัญหานั้น ๆ ตลอดจนความสามารถในการติดตามการพิสูจน์

จากแนวการเขียนข้อสอบวัดพฤติกรรมการเรียนรู้ด้านพุทธิพิสัย (Cognitive Domain) โดยทั่วไปจะประกอบด้วย การวัดพฤติกรรมด้านความรู้ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า สำหรับการเขียนข้อสอบวัดพฤติกรรมการเรียนรู้วิชาคณิตศาสตร์ก็มีลักษณะที่คล้ายคลึงกัน คือ วัดพฤติกรรมการเรียนรู้ด้านการคิดคำนวณ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ ในการศึกษาค้นคว้าวิจัยสนใจที่จะเขียนข้อสอบวัดพฤติกรรมการเรียนรู้ตามแบบของวิลสัน

4. ความเชื่อมั่นของแบบทดสอบ (Reliability)

4.1 ความหมายของความเชื่อมั่น

มีนักการศึกษา และนักจิตวิทยาหลายท่านได้ให้ความหมายของความเชื่อมั่นไว้ต่าง ๆ กัน ดังนี้

นันทาลลี (Nunnally. 1964 : 59) ให้ความหมายความเชื่อมั่นว่า เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการสอบ

ลินด์วอลล์และนิคโค (Lindvall and Nitko. 1967 : 126) ให้ความหมายความเชื่อมั่นว่าเป็นค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบสองครั้ง โดยใช้แบบทดสอบฉบับเดียวกันในเวลาต่างกัน

อนาสตาซี (Anastasi. 1968 : 71) ให้ความหมายความเชื่อมั่นว่าเป็นค่าคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันในเวลาที่แตกต่างกัน

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

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

ชวาล แพรัตกุล (2516 : 136) สำเริง บุญเรืองรัตน์ (2529 : 38) ให้ความหมายความเชื่อมั่นว่าเป็นความคงที่ของคะแนนที่ได้จากการทดสอบแบบทดสอบที่มีความเชื่อมั่นสูงไม่ว่าจะนำไปวัดกับบุคคลเดียวกันกี่ครั้งก็ตามก็จะได้ค่าคงที่เช่นเดิม ในทางตรงกันข้ามแบบทดสอบที่มีค่าความเชื่อมั่นต่ำ ค่าที่ได้จากการทดสอบสองครั้งของบุคคลกลุ่มเดียวกันจะมีการเปลี่ยนแปลงมาก

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

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

จากแนวคิดของนักการศึกษาและนักจิตวิทยาพอสรุปได้ว่าความเชื่อมั่นเป็นความคงที่ของคะแนนในการสอบด้วยแบบทดสอบฉบับเดียวกันหลาย ๆ ครั้งหรือจากแบบทดสอบคู่ขนานกับคนกลุ่มเดียวกัน

4.2 การประมาณค่าความเชื่อมั่น

เมห์เรนส์ และเลห์แมน (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 วิธีสัมประสิทธิ์อัลฟา (Coefficient Alpha)

4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt 's Analysis of Variance Procedure)

5. ความเชื่อมั่นของผู้ให้คะแนน (Score Judge Reliability)

อนันต์ ศรีโสภณ (2524 : 49 - 57) กล่าวถึงการคำนวณค่าความเชื่อมั่นว่ามีวิธีต่าง ๆ หลายวิธี ซึ่งในแต่ละวิธีทำให้เกิดความคลาดเคลื่อนไม่เหมือนกันการที่จะเลือกใช้วิธีใดนั้นขึ้นอยู่กับความมุ่งหมาย และลักษณะของคะแนนที่ได้จากการทดสอบนั้น วิธีหาค่าความเชื่อมั่นโดยทั่วไปมี 4 แบบ คือ

1. แบบสอบซ้ำ (Measures of Stability) เป็นการวัดความคงที่ของแบบทดสอบโดยนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มเดียวกันสองครั้งในเวลาที่แตกต่างกัน แล้วนำผลการทดสอบทั้งสอง มาหาค่าสัมประสิทธิ์สหสัมพันธ์โดยวิธี Pearson Product Moment Correlation การหาค่าความเชื่อมั่นโดยวิธีนี้มีความคลาดเคลื่อนที่เห็นได้ชัด คือ ช่วงระยะเวลาที่ทดสอบครั้งแรกและทดสอบครั้งหลังความแตกต่างของช่วงระยะเวลาทดสอบนี้นักเรียนจะอาศัยความจำจากการสอบครั้งแรกทำให้คะแนนสอบครั้งหลังสูงขึ้นและมีค่าความเชื่อมั่นสูง

2. แบบใช้ข้อสอบคล้ายกัน (Measures of Equivalence) เป็นการหาความเชื่อมั่นโดยใช้แบบทดสอบสองชุดที่มีเนื้อหา ค่าเฉลี่ย และค่าความแปรปรวนเท่ากันโดยทดสอบกับนักเรียนกลุ่มเดียวกันในวันเดียวกัน แล้วนำผลการสอบทั้งสองฉบับนั้นมาหาค่าสัมประสิทธิ์สหสัมพันธ์หรือค่าความเชื่อมั่น การหาค่าความเชื่อมั่นวิธีนี้มีความคลาดเคลื่อนเนื่องจากความแตกต่างของสเกล หรือความแตกต่างของข้อสอบสองฉบับ การที่จะสร้างให้มีค่าเฉลี่ยและความแปรปรวนเท่ากันนั้นทำได้ยากมากในทางปฏิบัติ

3. แบบใช้ข้อสอบคล้ายกัน และสอบซ้ำ (Measures of Equivalence and Stability) เป็นการหาความเชื่อมั่นโดยนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มเดียวกันในเวลาที่แตกต่างกัน

แล้วนำคะแนนผลการสอบทั้งสองมาหาค่าสัมประสิทธิ์สหสัมพันธ์ สาเหตุของการเปลี่ยนแปลงคะแนนจากวิธีนี้อาจเนื่องมาจากความไม่คงที่ของนักเรียน และตัวข้อสอบ ค่าความเชื่อมั่นที่ได้โดยวิธีนี้จะมีค่าต่ำกว่าสองวิธีแรก

4. แบบวัดความคงที่ภายใน (Measures of Internal Consistency) เป็นการหาความเชื่อมั่น โดยนำแบบทดสอบไปทดสอบกับนักเรียนเพียงครั้งเดียวแล้วนำไปคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแต่ละข้อกับคะแนนรวม การหาความเชื่อมั่นวิธีนี้แบ่งออกเป็น 4 แบบ คือ

4.1 แบบแบ่งครึ่งจำนวนข้อสอบ (Split - half) เป็นวิธีคำนวณค่าความเชื่อมั่นของแบบทดสอบโดยนำแบบทดสอบฉบับเดียวไปทดสอบกับนักเรียนกลุ่มหนึ่งครั้งเดียวแล้วแบ่งแบบทดสอบออกเป็นสองชุดนำคะแนนของทั้งสองชุดมาหาค่าสัมประสิทธิ์ r_{tt} คือค่าความเชื่อมั่นของแบบทดสอบฉบับย่อยที่แยกไป หรือเป็นค่าความเชื่อมั่นของแบบทดสอบครึ่งหนึ่งของแบบทดสอบฉบับเดิม ถ้าต้องการทราบค่าความเชื่อมั่นทั้งฉบับก็ปรับขยายด้วยสูตรของ Spearman Brown ค่าที่ได้คือค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4.2 Kuder - Richardson Estimates เป็นการหาความเชื่อมั่นโดยใช้สูตรของ 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 เป็นการหาความเชื่อมั่นของแบบทดสอบโดยการวิเคราะห์ค่าความแปรปรวน (Analysis of Variance) หรือเรียกว่า ANOVA วิธีนี้ให้ค่าความเชื่อมั่นเหมือนกับ KR - 20

บุญเชิด ภิญโญนนท์พงษ์ (2537 : 12 - 46) ได้สรุปแนวคิดในการหาความเชื่อมั่นของแบบทดสอบไว้ 2 แนวคิด คือ

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

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

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

1. แบบจำลองความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts) เป็นการประมาณค่าความเชื่อมั่นจากแบบทดสอบแต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิม ที่มีข้อตกลงอย่างเคร่งครัด 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 โดยผ่อนปรนเงื่อนไข ข้อ 2, 3 และ 4 ของแบบจำลองความคู่ขนานแบบมาตรฐานเดิมให้มีความเป็นไปได้มากขึ้น เนื่องจากในทางปฏิบัติแทบไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานตามแบบมาตรฐานเดิมได้ จึงได้มีการพัฒนาเทคนิคที่เหมาะสมขึ้นมาใหม่ ดังนี้

2.1 คะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากัน แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วน นั่นคือคะแนนจริงส่วนที่หนึ่งเท่ากับคะแนนจริงส่วนที่สองรวมกับค่าคงที่ค่าหนึ่ง ดังนี้ $T_{ig} = T_{ih} + C_{gh}$ เมื่อ $g = h = 1, \dots, k$ และ C_{gh} 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) เป็นการประมาณค่าความเชื่อมั่นที่ผ่อนปรนเงื่อนไขต่าง ๆ โดยคงไว้เฉพาะเงื่อนไขข้อ 1 ที่ว่าแต่ละส่วนของแบบทดสอบเนื้อหาต้องเป็นเอกพันธ์ หรือคุณลักษณะเดียวกัน (บุญเชิด ภิญโญนนตพงษ์. 2536. 12 ; อ้างอิงจาก Kristof. 1974 : 492)

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

การประมาณค่าความเชื่อมั่นนอกจากจะใช้สถิติขั้นสูงวิเคราะห์ด้วยเครื่องคอมพิวเตอร์แล้วนักทฤษฎีการวัดยังได้เสนอเทคนิคที่ใช้เครื่องคำนวณแบบธรรมดาไว้หลายเทคนิค ดังนี้

ฮอร์ส (บุญเชิด ภิญโญนนตพงษ์. 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

$$r_{tt} = \frac{[s_{x_1x_2}s_{x_1x_3} + s_{x_1x_2}s_{x_2x_3} + s_{x_1x_3}s_{x_2x_3}]^2}{s_x^2(s_{x_1x_2})(s_{x_1x_3})(s_{x_2x_3})}$$

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

เฟลด์ท (บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Feldt. 1975 : 557 - 561) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบที่แบ่งสามส่วนไม่เท่ากันของคริstofให้สามารถวิเคราะห์กับแบบทดสอบที่แบ่งสองส่วนไม่เท่ากันโดยเพิ่มข้อตกลงอีกหนึ่งข้อคือความแปรปรวนของคะแนนความคลาดเคลื่อนต้องเป็นไปตามทฤษฎีการวัดแบบมาตรฐานเดิมและไม่จำเป็นต้องทราบจำนวนข้อของแบบทดสอบ ดังนี้

$$r_{tt} = \frac{[s_{x_1x_2} / \lambda_1 \lambda_2]}{s_x^2}$$

เมื่อ $\lambda_1 = \frac{[s_{x_1}^2 + s_{x_1x_2}]}{s_x^2}$

$$\lambda_2 = \frac{[s_{x_2}^2 + s_{x_1x_2}]}{s_x^2}$$

r_{tt} แทน ค่าความเชื่อมั่นของแบบทดสอบ
 $\lambda_1 \lambda_2$ แทน สัดส่วนของจำนวนข้อแต่ละส่วนย่อย
 s_x^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ
 $s_{x_1}^2$ แทน ความแปรปรวนของคะแนนย่อยส่วนที่ 1
 $s_{x_2}^2$ แทน ความแปรปรวนของคะแนนย่อยส่วนที่ 2
 $s_{x_ix_j}$ แทน ความแปรปรวนร่วม

ราชู (บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Raju. 1977 : 549 - 565) ได้พัฒนาสูตรการประมาณค่าความเชื่อมั่นเมื่อแบ่งแบบทดสอบออกเป็นหลาย ๆ ส่วนแต่ละส่วน

มีความยาวไม่เท่ากัน 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 ภิญญอนันตพงษ์. 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 ภิญโญนันตพงษ์. 2537 : 33 - 37 ; อ้างอิงจาก Liou. 1989 : 153 - 163) ได้ทำการวิจัย และเสนอสัมประสิทธิ์ค่าความเชื่อมั่นขึ้นมาใหม่สองสัมประสิทธิ์ คือ r_{L1} และ r_{L2} ซึ่งมีข้อดีที่สามารถคำนวณด้วยเครื่องคำนวณธรรมดา และยังสามารถพิสูจน์ได้ทั้งเชิงทฤษฎี และจำนวนตัวเลขว่ามีค่าเท่ากับสัมประสิทธิ์ r_{LK} ของคริสทอฟ (Kristof) และ r_{LF} และของเฟลด์ท์ (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	ผลรวมของความแปรปรวนร่วมระหว่างส่วนย่อยของ แบบทดสอบ

บุญเชิด ภิญโญนนตพงษ์ (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

การแสดงผลมาของสัมประสิทธิ์ r_B

บุญเชิด ภิญโญนนตพงษ์ (2542 : 15 - 19) กล่าวถึงสัมประสิทธิ์ r_B ในการประมาณค่าความเชื่อมั่น ของแบบทดสอบแบบคะแนนจริงสัมพันธ์ แบบมาตรฐานเดิม ที่ได้พัฒนามาจากสัมประสิทธิ์ r_{12} ของเลียว (บุญเชิด ภิญโญนนตพงษ์. 2542 : 15 - 19 ; อ้างอิงจาก Liou. 1989) ที่มี K ข้อ ให้คะแนนแบบ (0 - 1) หรือตอบถูกให้ 1 ตอบผิด หรือไม่ตอบให้ 0 โดยแสดงที่มาของสัมประสิทธิ์ 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

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

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

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

จากแบบจำลองคะแนนของทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory Model) ได้คะแนนแต่ละข้อดังนี้

$$(4) \quad X_i = T_i + E_i \quad ; \quad i = 1 \text{ ถึง } K$$

และกำหนดข้อตกลงว่า แบบทดสอบประกอบด้วยข้อสอบที่วัดคุณลักษณะเดียวกัน และเป็นไปตามข้อตกลงแบบคะแนนจริงสัมพันธ์ คือ คะแนนจริงของแต่ละข้อมีความสัมพันธ์เป็นเส้นตรงกับคะแนนจริงทั้งฉบับ

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

เมื่อ λ_i เป็นลักษณะประจำของข้อที่ i ซึ่งกำหนดให้มีค่าเป็นจำนวนจริงบวก

C_i เป็นค่าคงที่ของข้อสอบ i

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

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

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

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

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

ผลรวมของความแปรปรวนของคะแนนสอบทุกข้อ คือ

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

นำ σ_x^2 ไปหารสมการ (9) และนิยามความเชื่อมั่นของแบบทดสอบ $\rho_{xx'} = \sigma_T^2 / \sigma_x^2$ จะได้สูตรความเชื่อมั่นซึ่งราซ (บุญเชิด ภิญโญนนตพงษ์. 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 ความแปรปรวนของคะแนนคลาดเคลื่อนแต่ละข้อ และ ทั้งฉบับจะมีความสัมพันธ์กันดังนี้ $\sigma_{Ei}^2 = \lambda_i \sigma_E^2$ และถ้านำคะแนนสอบของแต่ละข้อมาสร้างเป็นเมตริกซ์ความแปรปรวน ความแปรปรวนรวมแล้ว จะสามารถหาค่า λ_i ของแต่ละข้อจากแต่ละแถว (Row) ของเมตริกซ์ดังกล่าวได้โดยนำค่า σ_x^2 ไปหารผลรวมของความแปรปรวน ความแปรปรวนรวมในแต่ละแถว ดังนี้

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

จากสมการ (11) นำมาเขียนใหม่ โดยแยกความแปรปรวนรวมของคะแนนตัวมันเองออกมาซึ่งก็คือ ค่าความแปรปรวนของข้อนั้น เช่น

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

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

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

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

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

เมื่อ P เป็นสัดส่วนของผู้ตอบถูก หรือความยากง่ายของข้อสอบ

สำหรับสัมประสิทธิ์สหสัมพันธ์ จะต้องใช้สูตรกรณีเฉพาะของสัมประสิทธิ์สหสัมพันธ์ของ Pearson คือ สัมประสิทธิ์ ϕ ซึ่งมีสูตร ดังนี้

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

เมื่อ A, B, C และ D เป็นสัดส่วนจำนวนผู้ตอบข้อสอบสองข้อถูก หรือ ผิด ดังแสดงในตาราง

		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

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

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

ดังนั้นจากสมการ (13) และ (15) ได้ความแปรปรวนร่วม ดังนี้

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

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

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 ที่สร้างขึ้นใหม่จึงมีสูตร 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 อย่างเช่น ไฮเซ และบอร์นสเต็ด (Heise and Lohrnstedt) ได้เสนอสูตร Ω ซึ่งเป็นสัมประสิทธิ์ค่าความเชื่อมั่นของแบบทดสอบเมื่อกำหนดน้ำหนักความสำคัญเท่ากัน ต่อมา

อัลเลน (Allen) ได้พัฒนาสูตรของไฮช และบอร์นสเต็ดเมื่อพบว่าการกำหนดน้ำหนักความสำคัญไม่เท่ากันจะส่งผลต่อการประมาณค่าความเชื่อมั่น และความคลาดเคลื่อน พร้อมทั้งได้เสนอสูตร 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 โดยใช้สูตรคำนวณหาความสัมพันธ์สหสัมพันธ์ของเพียร์สัน (Pearson Product - Moment Coefficient Correlation) ถ้ามีค่าเหมือนเดิมแสดงว่ามีความคงที่ของคะแนน เป็นวิธีการทดสอบซ้ำ (Test Retest Method) การหาความเชื่อมั่นวิธีนี้มีจุดอ่อนตรงช่วงเวลาการสอบอาจทำให้การประมาณค่าความเชื่อมั่นของแบบทดสอบต่ำเกินไป หรือสูงเกินไป ซึ่งผลของการเปลี่ยนแปลงค่าความเชื่อมั่นนี้เรียกว่าเกิด Carry - Over Effect

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

สัมประสิทธิ์สหสัมพันธ์ โดยสูตรของเพียร์สัน การหาความเชื่อมั่นแบบนี้ มีปัญหาตรงการสร้างเครื่องมือวัดให้คู่ขนานกันเพราะเป็นเรื่องค่อนข้างยากที่จะสร้างเครื่องมือวัดให้คู่ขนานกันได้

1.3 ความเชื่อมั่นโดยใช้ความสอดคล้องภายใน (Internal Consistency Reliability) เป็นการหาความเชื่อมั่นที่ใช้แบบทดสอบฉบับเดียว ทำการทดสอบเพียงครั้งเดียว มีวิธีการประมาณค่าความเชื่อมั่น ดังนี้

1.3.1 วิธีแบ่งครึ่งแบบทดสอบ (Split - half Method) โดยวิเคราะห์ค่าความยากของแบบทดสอบเป็นรายข้อ แล้วเรียงข้อสอบจากข้อง่ายไปข้อยากนำไปทดสอบกับนักเรียน แล้วแบ่งแบบทดสอบเป็นสองส่วนโดยแบ่งให้แต่ละส่วนมีลักษณะเป็นคู่ขนานกันได้ฉบับข้อคู่กับฉบับข้อคี่หาค่าสหสัมพันธ์ระหว่างคะแนนโดยสูตรของเพียร์สันได้ค่าความเชื่อมั่นเพียงครึ่งฉบับ แล้วหาความเชื่อมั่นทั้งฉบับ โดยสูตรปรับขยายของสเปียร์แมน - บราวน์ (Spearman - Brown Formula)

1.3.2 วิธีคูเดอร์ - ริชาร์ดสัน (Kuder - Richardson Procedure) เครื่องมือที่จะหาค่าความเชื่อมั่นวิธีนี้ต้องมีลักษณะที่ข้อต้องประกอบร่วมกัน และคะแนนแต่ละข้อต้องอยู่ในลักษณะที่ทำถูกต้อง 1 คะแนน ทำผิดได้ 0 คะแนน วิธีนี้มีสูตรที่ใช้หาความเชื่อมั่นอยู่ 2 สูตร คือ สูตร KR - 20 กับสูตร KR - 21

1.3.3 วิธีของครอนบัค (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) ที่ครอนบัค เกลเซอร์ และราจารัตนัม (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 ในแบบทดสอบฉบับเดียวกันของผู้สอบคนเดียวกัน (Pretest - Posttest)

1.6 ความเชื่อมั่นจากการแบ่งส่วนย่อยของแบบทดสอบ ในบางครั้งแบบทดสอบที่สร้างจะแบ่งออกเป็นส่วนย่อย ๆ แต่ละส่วนย่อย ๆ (Parts) นั้น มีจำนวนข้อสอบไม่เท่ากัน หรือมีจำนวนข้อสอบเท่ากันแต่มีการกระจายคะแนนไม่เท่ากันและส่วนย่อย ๆ นั้นวัดเนื้อหาเดียวกัน การหาความเชื่อมั่นของแบบทดสอบที่เหมาะสม ควรหาในรูปของคอนเจนเนอริก (Congeneric) มีข้อตกลงว่าส่วนย่อยแต่ละส่วนนั้นมีคะแนนจริง (True Score) สัมพันธ์กันในเชิงเส้นตรงการหาความเชื่อมั่นแบบนี้ เป็นการหาความสอดคล้องภายใน โดยเอาคะแนนสอบเพียงครั้งเดียวของแบบทดสอบมาคำนวณ เมื่อแบบทดสอบแบ่งส่วนย่อยเป็นสองส่วน คำนวณโดยใช้สูตรของราจู (Raju. 1970.) แบบทดสอบแบ่งส่วนย่อยเป็นสามส่วน คำนวณโดยใช้สูตรของคริสทอฟ (Kristof. 1971.) แบบทดสอบแบ่งส่วนย่อยเป็นหลายส่วน คำนวณโดยใช้สูตรของราจู (Raju. 1977.) และเฟลด์ท์ (Feldt. 1975.)

อิทธิพลที่มีต่อค่าความเชื่อมั่นของแบบทดสอบที่มาจากการสอบครั้งเดียว

สูตรที่คำนวณหาความเชื่อมั่นของแบบทดสอบที่มาจากการสอบเพียงครั้งเดียวหรือคำนวณค่าความสอดคล้องภายในนั้นจำนวนข้อมีผลต่อค่าความเชื่อมั่น เพราะจำนวนข้อสอบมากข้อจะทำให้ค่าความแปรปรวนสูง การคำนวณค่าความเชื่อมั่นจะสูงด้วย และทำนองเดียวกันจำนวนข้อสอบน้อยจะทำให้ความแปรปรวนต่ำ การหาความเชื่อมั่นจะมีค่าต่ำด้วย ดังนั้นถ้าต้องการให้แบบทดสอบมีความเชื่อมั่นสูง ๆ จึงควรออกข้อสอบมาก ๆ ข้อ

2. ความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ หมายถึง ผลของคะแนนสอบที่สอบได้มีความคงที่ในการเป็นผู้รอบรู้ในเรื่องที่สอบ มีวิธีหาหลายวิธี ดังนี้

2.1 ความเชื่อมั่นแบบหาความคงที่ของความรอบรู้ (Stability Reliability) เป็นการคำนวณหาความเชื่อมั่นโดยการนำแบบทดสอบอิงเกณฑ์มาสอบซ้ำ 2 ครั้ง แล้วนำคะแนนที่ได้มาหาความคงที่ของการรอบรู้จากการกำหนดคะแนนจุดตัดที่เหมาะสม คำนวณโดยสูตรของชรอกและคอสแกร์ลี

2.2 ความเชื่อมั่นแบบความสอดคล้องในการตัดสินใจ (Decision Consistency Reliability) เป็นการหาความสอดคล้องระหว่างการทดสอบ 2 ครั้ง จากแบบทดสอบฉบับเดียวหรือแบบทดสอบที่คู่ขนานกัน 2 ฉบับ

2.3 การหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ จากการทดสอบเพียงครั้งเดียว มีหลายวิธี คือ

2.3.1 ความเชื่อมั่นจากสูตรของลิวิสตัน

2.3.2 ความเชื่อมั่นจากสูตรไบนอมิยัล (Binomial Formula)

2.3.3 ความเชื่อมั่นโดยใช้วิธีวิเคราะห์ความแปรปรวน

2.3.4 ความเชื่อมั่นโดยวิธีแบ่งครึ่งแบบทดสอบ

2.3.5 ความเชื่อมั่นโดยสูตรของแฮร์ริส

2.4 ความเชื่อมั่นที่คำนึงจุดประสงค์เชิงพฤติกรรมของแบบทดสอบแบบอิงเกณฑ์

การหาค่าความเชื่อมั่นของแบบทดสอบแบบอิงกลุ่มและอิงเกณฑ์มีหลายวิธี ขึ้นอยู่กับจุดประสงค์ในการสร้างแบบทดสอบ ชนิดของแบบทดสอบ ค่าความเชื่อมั่นของแบบทดสอบที่มีผลการวัดที่คงที่แน่นอนเชื่อถือได้ควรมีค่าสูงกว่า 0.70

4.3 องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบ

กรอนแลนด์ (Gronlund, 1976 : 117) ได้กล่าวถึงองค์ประกอบต่าง ๆ ที่มีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบไว้ ดังนี้

1. ความยาวของข้อสอบ (Length of Test) แบบทดสอบที่มีจำนวนข้อสอบมากจะมีความเชื่อมั่นสูงกว่าแบบทดสอบที่มีจำนวนข้อสอบน้อย

2. การกระจายของคะแนน (Spread of Score) ขึ้นอยู่กับความสามารถของกลุ่มผู้สอบ ถ้าความสามารถของผู้สอบในกลุ่มแตกต่างกันมาก การกระจายของคะแนนมีมากทำให้ค่าความเชื่อมั่นสูง แต่ถ้าความสามารถของผู้สอบในกลุ่มใกล้เคียงกันมากการกระจายของคะแนนมีน้อยทำให้ค่าความเชื่อมั่นต่ำ

3. ความยากของข้อสอบ (Difficulty of Test) แบบทดสอบที่ยาก หรือง่ายมากเกินไป จะทำให้ความแปรปรวนของคะแนนน้อยความเชื่อมั่นจะมีค่าต่ำ เพราะค่าความเชื่อมั่นของแบบทดสอบเกี่ยวข้องกับความแปรปรวนของคะแนนที่สอบได้

4. ความเป็นปรนัย (Objective) แบบทดสอบที่มีความเป็นปรนัยสูงจะมีค่าความเชื่อมั่นสูงด้วย

โกวิท ประวาลพุกษ์ และ สมศักดิ์ สินธุระเวชชัย (2523 : 177 - 179) ได้กล่าวถึงองค์ประกอบที่มีผลต่อค่าความเชื่อมั่น ดังนี้

1. ความยาวของแบบทดสอบ
2. ความเป็นเอกพันธ์ของกลุ่มนักเรียน
3. วิธีที่ใช้ในการคำนวณ
4. ความเป็นปรนัยในการให้คะแนน
5. ภาวะที่ไม่เป็นปกติในการสอบ
6. ความยากของแบบทดสอบ

ดังได้กล่าวมาข้างต้น การประมาณค่าความเชื่อมั่นของแบบทดสอบมีหลายวิธี ขึ้นอยู่กับจุดประสงค์ในการสร้างแบบทดสอบ และเนื่องจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ใช้ในการศึกษาค้นคว้าครั้งนี้ สร้างตามพฤติกรรมกรรมการเรียนรู้ของวิลสันที่มีการจัดเรียงข้อคำถามตามค่าความยากจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด ผู้วิจัยจึงสนใจที่จะประมาณ ค่าความเชื่อมั่นของแบบทดสอบจากสูตรสัมประสิทธิ์ r_B ที่เสนอโดย บุญเชิด ภิบุญญ อนันตพงษ์ (2538 : 51)

5. ทฤษฎีการตอบข้อสอบ (Item Response Theory)

บุญชม ศรีสะอาด (2540 : 204 - 205) กล่าวถึงทฤษฎีการตอบข้อสอบว่า มีชื่อเรียกอย่างอื่นอีก เช่น ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) หรือ ทฤษฎีโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve Theory) หรือ ทฤษฎีการทดสอบสมัยใหม่ (Modern Test Theory) ใช้ตัวย่อ IRT เป็นทฤษฎีที่ถือว่าคุณสมบัติหรือ คุณลักษณะต่าง ๆ ของข้อสอบ เช่น ค่าความยาก ค่าอำนาจจำแนก ค่าโอกาสในการเดาเป็นค่าที่ประจำอยู่กับข้อนั้นอย่างคงที่ ไม่ขึ้นอยู่กับความเก่งอ่อนของกลุ่มผู้สอบ (การไม่แปรไปตามกลุ่มตัวอย่าง ที่เรียกว่า Sample - Free) และถือว่าความสามารถของผู้สอบ (Ability) เป็นคุณสมบัติหรือคุณลักษณะที่ประจำอยู่ภายในตัวผู้สอบไม่ควรค่าเปลี่ยนไปตามค่าความยากของข้อสอบ (มีลักษณะเป็น Test - Free) นั่นคือ ไม่เอาคุณสมบัติของข้อสอบไปสัมพันธ์กับความสามารถของผู้สอบ และไม่เอาความสามารถของผู้สอบไปสัมพันธ์กับคุณสมบัติของข้อสอบ ทำให้เครื่องมือวัด และ สิ่งที่จะถูกวัดเป็นอิสระแก่กัน มีชื่อเรียกทฤษฎีเป็นภาษาไทยอีกหลายชื่อ เช่น ทฤษฎีการตอบสนองข้อสอบ ทฤษฎีการสนองตอบข้อสอบ ทฤษฎีการตอบข้อกระทง ทฤษฎีคุณลักษณะภายใน ฯลฯ

ทฤษฎีการตอบข้อสอบกล่าวว่าผลการสอบของผู้สอบจากแบบทดสอบใด ๆ นั้นขึ้นอยู่กับความสามารถของผู้สอบ เขียนเป็นสมการได้ว่า (สำเร็จ บุญเรืองรัตน์. 2525 : 3)

$$P = f(\theta)$$

เมื่อ	P	แทน	ผลการสอบ (Performance)
	θ	แทน	ความสามารถ (Ability หรือ Trait)
	f	แทน	ความสัมพันธ์ (Function)

มีรูปแบบแสดงความสัมพันธ์ระหว่างผลการสอบกับความสามารถของผู้สอบที่ละเอียดลงไปหลายรูปแบบ โดยเขียนในสมการที่แสดงถึงโอกาสที่ผู้สอบที่มีระดับความสามารถ θ จะตอบคำถามข้อใด ๆ ได้ถูกต้อง เช่น รูปแบบที่เสนอโดยลอร์ด (F.M. Lord) เรียกว่า Normal Ogive Model รูปแบบที่เสนอโดยราซช์ (Rasch) เรียกว่า One Parameter Logistic Model (อาจใช้ตัวย่อว่า 1 - P หรือ 1PL) หรือ บางครั้งเรียกว่า Rasch Model ซึ่งถือข้อตกลงว่าข้อสอบ

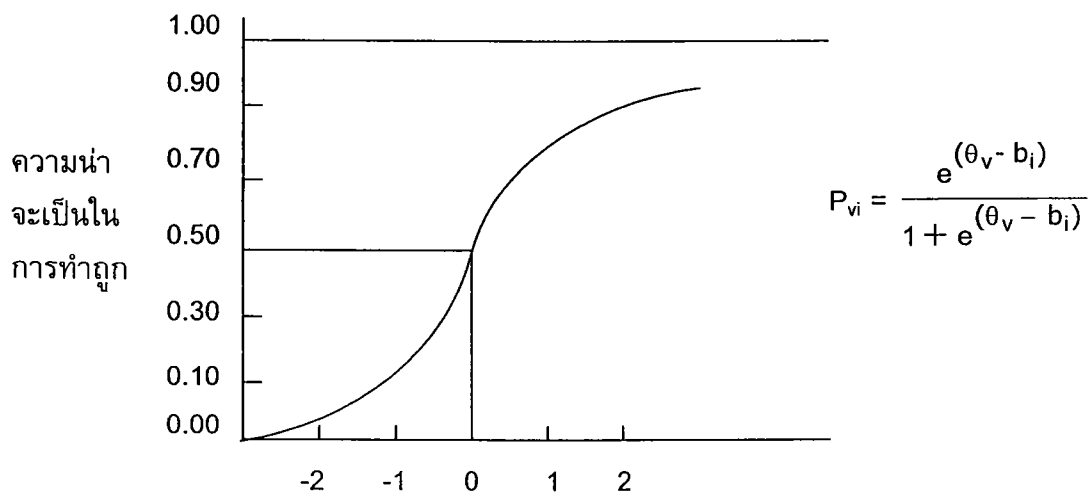
ทุกข้อมีค่าอำนาจจำแนก (a) เท่ากันหมด รูปแบบที่เสนอโดยเบิร์นบอม (A. Birnbaum) เรียกว่า Two - Parameter Logistic Model (อาจใช้ตัวย่อว่า 2 - P หรือ 2PL) ในสมการจะมีค่าความยาก (b) และอำนาจจำแนก (a) อยู่ด้วย และ Three - Parameter Logistic Model (อาจใช้ตัวย่อว่า 3 - P หรือ 3PL) ในสมการจะเพิ่มค่าโอกาสของการเดา (C) เข้ามาจึงมี Parameter สามตัว

ราชคโมเดล หรือ โมเดลที่ใช้พารามิเตอร์ 1 ตัว (One Parameter Logistic Model) เป็นโมเดลที่ใช้ในการหาค่าความยากของข้อสอบที่ไม่ต้องไปสัมพันธ์กับคน และหาค่าความสามารถของคนโดยไม่ต้องสัมพันธ์กับระดับความยากง่ายของข้อสอบ เป็นการวัดคุณลักษณะของคนและข้อสอบ โดยการกำหนดให้ θ_v เป็นพารามิเตอร์ที่แทนความสามารถของคนที่ v และ b_i เป็นพารามิเตอร์ที่แทนค่าความยากของข้อสอบข้อที่ i พารามิเตอร์ทั้งสองจึงแทนตำแหน่งของคนและตำแหน่งของข้อสอบในฐานะเป็นตัวแปรแฝงซึ่งกันและกัน เป็นโมเดลที่ใช้ประมาณค่าความน่าจะเป็นที่บุคคล v จะทำข้อสอบข้อที่ i ได้ถูกต้อง (Rasch. 1950 : 62 – 125)

ในการทำข้อสอบข้อหนึ่ง ควรจะใช้ความสามารถต่อสู้กับความยากของข้อสอบค่าพารามิเตอร์ θ_v และ b_i จึงเกี่ยวข้องกันในรูปแบบของ $(\theta_v - b_i)$ ซึ่งผลต่างนี้เองจะเป็นตัวบังคับค่าความน่าจะเป็นที่คาดว่าจะเกิดขึ้นเมื่อบุคคลใช้ความสามารถต่อสู้กับความยากของข้อสอบ i ค่าพารามิเตอร์แต่ละตัวจะแปรผันในช่วง $-\infty$ ถึง $+\infty$ ดังนั้นค่าของ $(\theta_v - b_i)$ จึงมีค่าแปรผันอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ด้วย แต่ค่าความน่าจะเป็นใด ๆ ก็ตาม จะมีค่าอยู่ระหว่าง 0 ถึง 1 เสมอ ดังนั้นค่าของ $(\theta_v - b_i)$ จึงถูกแปลงให้อยู่ในรูปของเอกซ์โปเนนเชียลฐาน e คือ $e^{(\theta_v - b_i)}$ และถ้ากำหนดให้ P_{vi} แทนค่าความน่าจะเป็นที่บุคคล v จะทำข้อสอบ i ได้ถูกต้องแล้ว จะได้สมการสำหรับวิเคราะห์ความน่าจะเป็นในการทำข้อสอบที่เรียกว่า ราชคโมเดล (Rasch Model) ดังนี้

$$P_{vi} = \frac{e^{(\theta_v - b_i)}}{1 + e^{(\theta_v - b_i)}} \quad \dots\dots\dots(1)$$

และจะทำให้ได้เส้นความถี่สะสมปกติหรือ Normal Ogive ที่ราชคเรียกว่า Logistic ที่แสดงได้ดังรูป



ภาพแสดงความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ

จากรูปเมื่อบุคคล v มีความสามารถมากกว่าความยากของข้อสอบ i ค่า θ_v จะมากกว่าค่า b_i และค่าของ $(\theta_v - b_i)$ จะเป็นบวก ซึ่งแสดงว่า ความน่าจะเป็นที่บุคคล v ทำข้อสอบถูกมีมากกว่า 0.5 ถ้าผลต่างนี้มีค่าเป็นบวกและเพิ่มขึ้น แสดงว่าความน่าจะเป็นในการทำข้อสอบถูกของบุคคล จะเพิ่มขึ้นด้วย แต่ถ้าเมื่อใดที่ข้อสอบยากเกินไปสำหรับบุคคล ค่า θ_v จะน้อยกว่า b_i และทำให้ $(\theta_v - b_i)$ เป็นค่าลบ ซึ่งแสดงว่าความน่าจะเป็นในการทำข้อสอบถูกมีน้อยกว่า 0.5

ถ้ากำหนดให้บุคคล v ทำข้อสอบ ข้อ i ถูกต้องจะได้คะแนน $X_{vi} = i$ และเมื่อบุคคล v ทำข้อสอบ i ผิด จะได้คะแนน $X_{vi} = 0$ แล้ว สมการของราชค์ จะกลายเป็น

$$\text{pr} \{ X_{vi} / \theta_v, b_i \} = e^{X_{vi}(\theta_v - b_i)} / (1 + e^{(\theta_v - b_i)}) \quad \dots\dots\dots (2)$$

ถ้าบุคคล v แปรค่าจาก 1 ถึง N และข้อสอบ i แปรค่าจาก 1 ถึง L แล้ว คะแนนที่คนสอบได้จะเป็น

$$r_v = \sum_{i=1}^L x_{vi} \quad \text{เมื่อ } v = 1 - N$$

และจำนวนคนที่ทำข้อสอบถูกในแต่ละข้อจะเป็น

$$s_i = \sum_{v=1}^N x_{vi} \quad \text{เมื่อ } i = 1 - L$$

เมื่อคำนวณความน่าจะเป็นของการที่ทุกคนทำข้อสอบทุกข้อแล้ว ตัวแปรที่จำเป็น ได้แก่ ความสามารถของทุกคนและความยากของข้อสอบทุกข้อ ซึ่งจะเป็นตัวควบคุมคะแนน ทั้งหมดและถ้านำค่าเหล่านี้ไปแทนค่าในสมการข้างต้น จะพบว่า ความน่าจะเป็นทั้งหมดเป็น ผลคูณของความสามารถของคนกับความยากของข้อสอบนั่นเอง ดังสมการ

$$p\{(x_{vi}) / (\theta_v), (b_i)\} = \sum_1^N \sum_1^L \left[\frac{e^{(X_{vi}(\theta_v - b_i))}}{1 + e^{(\theta_v - b_i)}} \right] \dots\dots\dots (3)$$

แต่เพราะว่า

$$\sum_v^N \sum_1^L x_{vi} \theta_v = \sum_v^N r_v \theta_v$$

$$\text{และ} \quad \sum_v^N \sum_1^L x_{vi} b_i = \sum_1^L s_i b_i$$

แทนค่าในสมการ 3 จะได้

$$\begin{aligned} p\{(x_{vi}) / (\theta_v), (b_i)\} &= \frac{e^{\left[\sum_v^N r_v \theta_v - \sum_1^L s_i b_i \right]}}{\sum_v^N \sum_1^L \left[1 + e^{\theta_v - b_i} \right]} \\ &= \frac{\left[e^{\left(\sum_v^N r_v \theta_v \right)} \right] \left[e^{-\left(\sum_1^L s_i b_i \right)} \right]}{\sum_v^N \sum_1^L \left[1 + e^{\theta_v - b_i} \right]} \end{aligned}$$

จากสมการข้างต้น จะเห็นว่าเราสามารถแยกคะแนนของคนออกจากคะแนนของ ข้อสอบ ทำให้สามารถวิเคราะห์ค่าความสามารถของคนโดยไม่ขึ้นกับข้อสอบ (Person Free) ได้ ในทำนองเดียวกันก็อาจวิเคราะห์ค่าของข้อคำถามโดยไม่ต้องขึ้นอยู่กับุคคล (Item Free) ได้ ซึ่งเป็นลักษณะพิเศษของการวิเคราะห์แบบราซช์ โมเดล ที่ทำให้สามารถหาค่าวิเคราะห์ของคน และข้อสอบได้เป็นอิสระจากกัน

โมเดลของราซค์ที่ใช้ในการคำนวณค่าความยากและค่าความสามารถนั้น เป็นโมเดลโลจิสติกที่ใช้พารามิเตอร์ตัวเดียว (One -- Parameter Logistic Model) โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด และไม่มีการเดาได้ถูกต้อง มีสมการดังนี้

$$P_g(\theta) = \frac{e^{D\bar{a}(\theta-b_g)}}{1 + e^{D\bar{a}(\theta-b_g)}} \quad (g = 1, 2, \dots, n)$$

เมื่อ $P_g(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อ g ได้ถูกต้อง

b_g แทน ค่าความยากของข้อสอบข้อ g

$\bar{a}$ แทน ค่าอำนาจจำแนกที่เป็นกลาง

D แทน ค่าคงที่ในการปรับมีค่าเท่ากับ 1.7

ต่อมา Wright และ Stone ได้ปรับสมการดังกล่าวเป็น

$$P_g(\theta) = \frac{e^{(\theta-b_g)}}{1 + e^{(\theta-b_g)}}$$

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยวิเคราะห์ค่าความยากของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามวิธีของราซค์ (Rasch Model) เพื่อใช้ในการจัดเรียงข้อสอบจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด

6. งานวิจัยที่เกี่ยวข้องกับการตรวจให้คะแนน

งานวิจัยในต่างประเทศ

มีผู้ศึกษาค้นคว้าเกี่ยวกับการตรวจให้คะแนนแบบต่าง ๆ ไว้ ดังนี้

คูมบ์ส และโวมเมอร์ (Coombs, Milholland and Womer. 1956 : 13 - 37) นำแนวคิดที่คูมบ์สเสนอไว้ มาทำการทดลองศึกษาเปรียบเทียบกับวิธีธรรมดา (0 - 1) โดยใช้แบบทดสอบ 3 ฉบับ ๆ ละ 40 ข้อ ได้แก่แบบทดสอบเกี่ยวกับคำศัพท์ ความรู้ของคนขับรถ และแบบทดสอบสายตาในด้านมิติสัมพันธ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นปีที่ 3 และ 4 ของโรงเรียนมัธยมแจคสัน (Jackson High School) รัฐมิชิแกน จำนวน 855 คน แบ่งออกเป็น 3 กลุ่ม แต่ละกลุ่มมีเงื่อนไขการตอบ และการตรวจให้คะแนนที่แตกต่างกัน คือวิธีของคูมบ์ส วิธี 0 - 1 และวิธีทั้งสองร่วมกัน

ผลการศึกษาพบว่าค่าความเชื่อมั่นจากการตรวจให้คะแนนวิธีของคัมป์สูงกว่าแบบ 0 - 1 ความเที่ยงตรงไม่ต่างกัน วิธีการตรวจให้คะแนนของคัมป์ป้องกันการเดาได้ดีกว่า และทัศนคติของผู้สอบชอบวิธีการตรวจของคัมป์มากกว่า

เดวิส และฟิฟเฟอร์ (Davis and Fiffer. 1959 : 159 - 169) ได้ศึกษาเกี่ยวกับค่าความเชื่อมั่น ค่าความเที่ยงตรงของแบบทดสอบวัดความถนัด และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยกำหนดน้ำหนักของคะแนนแต่ละตัวเลือกไม่เท่ากันโดยผู้เชี่ยวชาญเปรียบเทียบกับวิธีการตรวจให้คะแนนแบบ 0 - 1 เพราะคิดว่าการคิดโจทย์คณิตศาสตร์ปัญหานั้นตัวเลือกแต่ละตัวเลือกมีความยากง่ายไม่เท่ากัน กลุ่มตัวอย่างที่ศึกษาเป็นนักเรียนการบินที่ Lackland Air Force Base จำนวน 65 คน ผลการศึกษาพบว่าค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนด้วยวิธีให้น้ำหนักคะแนนทุกตัวเลือกสูงกว่าวิธีตรวจให้คะแนนแบบ 0 - 1 และมีความเที่ยงตรงสูงกว่า

อัลเบิร์ต (Albert. 1970 : 1619 - A) ได้ศึกษาวิธีการตรวจให้คะแนนความรู้บางส่วนของวิชาคณิตศาสตร์โดยใช้โจทย์คณิตศาสตร์จำนวน 25 ปัญหา กลุ่มตัวอย่างที่ศึกษามีจำนวน 2 กลุ่ม ๆ ละ 146 คน เป็นนักเรียนระดับเกรด 11 และวิธีการให้คะแนน 3 ลักษณะ ดังนี้ 1) แบบทดสอบเลือกตอบตัวเดียวมีวิธีการให้คะแนน 3 วิธี คือ วิธี 0 - 1 วิธีกำหนดน้ำหนักตัวเลือกโดยกำหนดจากโอกาสผิดพลาดมาก หรือน้อย และวิธีกำหนดน้ำหนักตัวเลือกโดยครูเป็นผู้กำหนด และ 2) แบบขยายคำตอบเป็นแบบที่มีการแก้ปัญหาหลายขั้นตอนมีวิธีการให้คะแนน 2 วิธี คือ ให้คะแนนข้อละ 1 คะแนนในการตอบที่ถูกต้อง และให้คะแนนขั้นละ 1 คะแนน ถ้านักเรียนเริ่มต้นจากข้อมูลที่ต้องการถึงคำตอบที่ต้องการ ผลการศึกษาแบบขยายคำตอบให้ค่าอำนาจจำแนกค่าความเชื่อมั่น และค่าความเที่ยงตรงกว่าแบบทดสอบเลือกตอบที่มีตัวเลือกถูกต้องตัวเดียว

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

อาบู ซาอีฟ และไดมอนด์ (Abu - Sayf and Diamond. 1976 : 62 - 65) ได้ศึกษาเกี่ยวกับค่าความเชื่อมั่น และค่าความเที่ยงตรงของแบบทดสอบที่มีวิธีการตรวจให้คะแนน 3 วิธี

คือ ให้คะแนนเฉพาะข้อที่ถูกและตอบอย่างมั่นใจ ให้คะแนนข้อที่ถูกโดยไม่ได้เดา และให้คะแนนทุกข้อที่ตอบถูกโดยไม่คำนึงถึงว่ามีความมั่นใจในการตอบถูกหรือไม่เพียงใดโดยใช้สูตรค่าความเชื่อมั่นครอนบักแอลฟา (Cronbach's Alpha) ผลการศึกษาพบว่าค่าความเชื่อมั่นในการตอบสูงสุดเท่ากับ .898 คือการให้คะแนนเฉพาะข้อที่ถูกและมีความมั่นใจในการตอบส่วนการให้คะแนนทุกข้อที่ตอบถูกโดยไม่ได้เดา และการให้คะแนนข้อที่ตอบถูกโดยไม่คำนึงถึงความมั่นใจ มีความเชื่อมั่น .871 และ .774 ตามลำดับ สำหรับค่าความเที่ยงตรงวิธีให้คะแนนเฉพาะข้อที่ถูกและมั่นใจ ให้คะแนนข้อที่ถูกโดยไม่ได้เดา และวิธีให้คะแนนทุกข้อที่ถูกโดยไม่คำนึงถึงความมั่นใจในการตอบมีค่าความเที่ยงตรง .260, .236 และ .219 ตามลำดับ

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

สมบุญ สूरียวงค์ (Suriyawongse, Somboon. 1987 : 613- A) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์ (ITBS) Form 8 จำนวน 3 ฉบับ คือ ฉบับที่ 1 ความเข้าใจ จำนวน 44 ข้อ ฉบับที่ 2 การแก้โจทย์ปัญหา จำนวน 32 ข้อ และฉบับที่ 3 ความรู้ด้านการคิดคำนวณจำนวน 45 ข้อที่มีการจัดเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดและตรวจให้คะแนนโดยใช้กฎชิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด และนับคะแนนตั้งแต่ข้อสอบข้อแรกถึงข้อที่ตอบถูกก่อนข้อที่ตอบผิดตำแหน่งนั้น ๆ ข้อสอบข้อที่ตอบถูกหลังข้อที่ตอบผิดจะไม่นำมานับคะแนนด้วยถือว่าเป็นข้อที่ตอบถูกเพราะการเดา โดยศึกษากับนักเรียนเกรด 8 ที่ Urban School District ในอเมริกาตะวันตกเฉียงใต้ (Southwestern United States) จำนวน 700 คน ซึ่งแบ่งเป็น 7 กลุ่ม กลุ่มละ 100 คน ผลการศึกษาพบว่าการตอบผิดตำแหน่งที่สองถึงห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่งและการไม่กำหนดตำแหน่งการตอบผิด โดยการตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด

วิลสัน และคณะ (Wilson and others. 1988 : 213 - 217) ได้ศึกษาความเชื่อมั่นของแบบทดสอบมาตรฐานวิชาคณิตศาสตร์ (ITBS) Form 8 จำนวน 1 ฉบับ คือ ความเข้าใจ จำนวน 44 ข้อที่มีการจัดเรียงข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดและตรวจให้คะแนนโดยใช้กฎชิลลิงจากการกำหนดตำแหน่งการตอบผิดต่างกัน 5 ตำแหน่ง กับการไม่กำหนดตำแหน่ง

การตอบผิด และนับคะแนนตั้งแต่ข้อสอบข้อแรกถึงข้อที่ตอบถูกต้องก่อนข้อที่ตอบผิดตำแหน่งนั้น ๆ ข้อสอบข้อที่ตอบถูกต้องหลังข้อที่ตอบผิดจะไม่นำมานับคะแนน ด้วยถือว่าเป็นข้อที่ตอบถูกต้องเพราะการเดา โดยศึกษากับนักเรียนเกรด 8 จำนวน 700 คน แบ่งเป็น 7 กลุ่ม ๆ ละ 100 คน ที่ Urban School District ในอเมริกาตะวันตกเฉียงใต้ (Southwestern United States) ผลการศึกษาพบว่าการตอบผิดตำแหน่งที่สองถึงห้าให้ค่าความเชื่อมั่นสูงกว่าการตอบผิดตำแหน่งที่หนึ่ง และการไม่กำหนดตำแหน่งการตอบผิด โดยการตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด

งานวิจัยภายในประเทศ

มีผู้ศึกษาค้นคว้าวิจัยเกี่ยวกับวิธีการตรวจให้คะแนนแบบต่าง ๆ ดังนี้

อรรพรรณ ดัณฑ์เจริญรัตน์ (2517 : 1 - 67) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีวิธีการตอบและวิธีการตรวจให้คะแนนแบบ 0 - 1 วิธีของคูมบัส และวิธีใหม่ตามแนวคิดของอนันต์ ศรีโสภณ กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาระดับประกาศนียบัตรการศึกษาตอนต้น วิทยาลัยครูหมู่บ้านจอมบึง จังหวัดราชบุรี จำนวน 214 คน สุ่มเป็น 3 กลุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวิชาภาษาไทยชนิด 5 ตัวเลือก จำนวน 3 ฉบับ ผลการศึกษาพบว่าแบบทดสอบที่มีวิธีตอบและวิธีตรวจโดยวิธีใหม่ตามแนวคิดของอนันต์ มีค่าความเชื่อมั่นของแบบทดสอบสูงกว่าวิธีอื่น ๆ

ธีรศักดิ์ อินทรมาตย์ (2520 : 63 - 73) ได้ศึกษาผลการตรวจให้คะแนน 3 วิธี คือ วิธีตอบแบบธรรมดา (0 - 1) วิธีบอกความมั่นใจในการตอบ วิธีตอบทุกตัวเลือกต่อค่าความเชื่อมั่น ค่าความเที่ยงตรง และปริมาณการเดาของแบบทดสอบวิชาวิทยาศาสตร์ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนสมุทรสงคราม จำนวน 216 คน โดยแบ่งเป็น 3 กลุ่ม แต่ละกลุ่มได้รับแบบทดสอบที่มีวิธีการตอบ และตรวจให้คะแนนแต่ละวิธีตามลำดับ ซึ่งพบว่าวิธีการตอบโดยบอกความมั่นใจในการตอบ มีความเชื่อมั่นสูงกว่าวิธีตอบแบบธรรมดาและวิธีตอบทุกตัวเลือก ปริมาณการเดาของนักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ในระดับต่ำจะตอบแบบทดสอบ โดยการเดามากกว่านักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์สูงกว่าในทุกวิธีการตอบ แบบทดสอบที่มีวิธีตอบทุกตัวเลือกมีค่าความเที่ยงตรงสูงสุด วิธีตอบแบบธรรมดาและแบบทดสอบที่มีการตอบโดยบอกความมั่นใจมีความเที่ยงตรงลดลงตามลำดับ แต่ทั้งสามวิธีมีค่าความเที่ยงตรงแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เพ็ญศรี สว่างเนตร (2520 : บทคัดย่อ) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบชนิดเลือกตอบด้วยเทคนิคการให้คะแนนต่างกัน 4 วิธี คือ วิธีกำหนดคะแนนในข้อที่ตอบถูกข้อละหนึ่งคะแนน (0 - 1) วิธีลดคะแนนข้อที่ตอบผิด วิธีเพิ่มคะแนนข้อที่เว้น และ วิธีทดสอบความมั่นใจ กลุ่มตัวอย่างที่ศึกษา เป็นนักศึกษาระดับประกาศนียบัตรวิชาการศึกษาชั้นสูงที่เรียนวิชาภาษาอังกฤษเป็นวิชาเอก หรือวิชาโท จำนวน 218 คน และหาค่าความเชื่อมั่นจากสูตรการวิเคราะห์

ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance) ซึ่งพบว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้เทคนิคให้คะแนนความมั่นใจ (.61) สูงกว่าแบบ 0 - 1 (.55) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเชื่อมั่นของแบบทดสอบที่ใช้เทคนิคลดคะแนนข้อตอบผิด (.56) มีค่าเพิ่มจากแบบ 0 - 1 เล็กน้อย และแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 จึงสรุปได้ว่าเทคนิคให้คะแนนความมั่นใจที่ใช้กับแบบทดสอบเลือกตอบมีค่าความเชื่อมั่นสูงที่สุด

สำราญ มีแจ้ง (2525 : 25 - 61) ได้ศึกษาผลของคำสั่ง และวิธีการให้คะแนนที่ต่างกัน ต่อค่าความเชื่อมั่น ค่าความเที่ยงตรง และ ค่าอำนาจจำแนกของแบบทดสอบเลือกตอบ ที่มีตัวเลือกถูกตัวเดียวโดยมีวิธีการให้คะแนน 4 วิธีคือ วิธี 0 - 1 วิธีคัมป์ส วิธีของอนันต์ และวิธีของสำราญกับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนสุรศักดิ์มนตรี และโรงเรียนวัดน้อยในที่เรียนวิชาคณิตศาสตร์ ค. 011 ปีการศึกษา 2524 จำนวน 240 คน โดยเมื่อผู้สอบทำเครื่องหมายตรงกับตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียวแล้วถูกต้องจะได้ 4 คะแนน ผลการศึกษาพบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและวิธีการตรวจให้คะแนนแบบสำราญสูงกว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0 - 1 ($P < .01$) และค่าความเชื่อมั่นที่ใช้คำสั่ง และการให้คะแนนตามวิธีของคัมป์สสูงกว่าค่าความเชื่อมั่นของแบบทดสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0 - 1 ($P < .05$)

เกริกชัย ฮวบเจริญ (2525 : 1 - 41) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นและคะแนนการเดาของการตอบ และตรวจให้คะแนนแบบทดสอบปรนัยเลือกตอบ ที่มีวิธีการตอบ 3 วิธี คือ วิธีธรรมดา วิธีทดสอบความมั่นใจ และวิธีใหม่ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2524 โรงเรียนวัดอ่างทอง จำนวน 114 คน แบ่งเป็น 3 กลุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวิชาวิทยาศาสตร์แบบปรนัยเลือกตอบชนิด 5 ตัวเลือก จำนวน 40 ข้อ ผลการศึกษาพบว่าวิธีตอบ และวิธีตรวจให้คะแนนใหม่ และวิธีทดสอบความมั่นใจ ให้ค่าความเชื่อมั่นสูงกว่าแบบธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และให้ค่าคะแนนการเดาของนักเรียนน้อยกว่าวิธีธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ทวี ทองคำ (2526 : บทคัดย่อ) ศึกษาเปรียบเทียบค่าความเชื่อมั่น ค่าความเที่ยงตรง และอำนาจจำแนกของแบบทดสอบวิชาคณิตศาสตร์พาณิชยกรรม 2 ชนิด 5 ตัวเลือก ที่ใช้คำสั่ง และวิธีการตรวจให้คะแนนที่ต่างกัน 3 วิธี คือ วิธี 0 - 1 วิธีของสำราญ วิธีของกิบบอนส์และคณะ กับนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ของวิทยาลัยเทคนิคยะลา 180 คน แบ่งเป็น 3 กลุ่มแต่ละกลุ่มได้รับแบบทดสอบที่มีวิธีการตอบและตรวจให้คะแนนแต่ละวิธี ซึ่งผลการศึกษาพบว่า ค่าความเชื่อมั่นของแบบทดสอบในทุกคำสั่งและวิธีการให้คะแนนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

นภา กาญจนกิจโสภณ (2528 : 21) ศึกษาเปรียบเทียบคะแนนสอบระหว่างการทำให้นักคะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิม กับการให้นักคะแนนรายข้อต่างกัน

ตามทฤษฎีการสนองตอบข้อคำถาม กับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 ที่มีอยู่ แล้วของสำนักทดสอบทางการศึกษากรมวิชาการ 2.7% ได้จำนวนทั้งสิ้นวิชาละ 3,907 คน ซึ่งข้อมูลเดิมตรวจให้คะแนนแบบให้น้ำหนักคะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิมจึงนำมาตรวจให้คะแนนแบบให้น้ำหนักรายข้อต่างกันตามทฤษฎีการสนองตอบข้อคำถาม โมเดลที่มีพารามิเตอร์ 3 ตัว ผลการศึกษาพบว่า การให้น้ำหนักคะแนนรายข้อต่างกันตามทฤษฎีการสนองตอบข้อคำถามแบบทดสอบวิชาคณิตศาสตร์ วิทยาศาสตร์และพลานามัยมีค่าความเชื่อมั่น .991, .994 และ .926 ตามลำดับ สูงกว่าจากการให้น้ำหนักคะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิมที่มีค่า .884 , .926 และ .437

พนิชา สวัสดิ์มงคล (2530 : 63 - 67) ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบสอบเลือกตอบที่วัดระดับความรู้ที่ต่างกัน 2 ระดับ คือระดับความรู้ขั้นต่ำและความรู้ขั้นสูง โดยมีวิธีการให้คะแนนวิธีธรรมดา กับวิธีการให้คะแนนความมั่นใจ ในกลุ่มผู้สอบที่มีผลสัมฤทธิ์ทางการเรียนสูง ปานกลางและต่ำ โดยศึกษากับนักศึกษาพยาบาลชั้นปีที่ 3 ปีการศึกษา 2530 หลักสูตรพยาบาลศาสตรบัณฑิต (วิทยาศาสตร์บัณฑิตสาขาพยาบาล) สังกัดทบวงมหาวิทยาลัย และองค์การการกุศลในเขตกรุงเทพมหานคร จำนวนที่เข้ารับการทดสอบจริง 195 คน เครื่องมือที่ใช้คือแบบทดสอบวัดผลสัมฤทธิ์ทางการพยาบาลผู้ป่วยระบบทางเดินหายใจ จำนวน 60 ข้อ เป็นข้อสอบที่วัดระดับความรู้ขั้นต่ำ 30 ข้อ วัดระดับความรู้ขั้นสูง 30 ข้อ ผลการศึกษาพบว่า 1) ค่าความเชื่อมั่นของแบบสอบวัดที่ระดับความรู้ขั้นต่ำ และมีวิธีให้คะแนนแบบธรรมดา ระหว่างกลุ่มต่ำ กับกลุ่มปานกลางและสูง มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 2) ค่าความเชื่อมั่นของแบบสอบวัดที่วัดระดับความรู้ขั้นสูง และมีวิธีให้คะแนนแบบธรรมดาที่ได้จากทั้ง 3 กลุ่มมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 3) ค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นต่ำและมีวิธีให้คะแนนความมั่นใจที่ได้จาก 3 กลุ่ม มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 4) ค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นสูงและมีวิธีให้คะแนนความมั่นใจระหว่างกลุ่มสูงกับกลุ่มต่ำมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผลจากการเปรียบเทียบค่าความเชื่อมั่นของแบบสอบที่วัดระดับความรู้ขั้นต่ำและแบบสอบที่วัดระดับความรู้ขั้นสูงที่ได้จากวิธีให้คะแนนทั้ง 2 วิธี มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทั้งในกลุ่มสูง ปานกลางและต่ำ

สุพัฒน์ สุขมลสันต์ (2528 : 47 - 60) ศึกษาเปรียบเทียบการให้น้ำหนักคะแนนที่ต่างกัน 5 วิธี คือ 1) วิธีใช้ค่าเดลต้า (Δ) 2) วิธีใช้ค่าเบต้าแปลง (β) 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 ผลการศึกษา

พบว่าความเชื่อมั่นของแบบทดสอบอิงกลุ่ม ค่าลำดับของแบบทดสอบทั้ง 2 ฉบับ ที่ให้คะแนนวิธีต่าง ๆ มีค่าสูงทุกวิธี โดยแบบทดสอบที่ให้คะแนนวิธีใช้น้ำหนักคะแนนที่เหมาะสมมีค่าความเชื่อมั่นสูงสุด

อนันต์ ศรีโสภา (2532 : 1 - 54) ทำการทดลองแบบใหม่ของข้อสอบแบบเลือกตอบ โดยสุ่มนักศึกษาจำนวน 144 คน ซึ่งลงทะเบียนเรียนวิชาวิจัยเบื้องต้น และ แบ่งกลุ่มทดลองเป็น 3 กลุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์แบบเลือกตอบวิชาวิจัยเบื้องต้นจำนวน 3 ฉบับ กลุ่มตัวอย่างที่ 1 ให้ตอบตามรูปแบบวิธีการตอบแบบใหม่ กลุ่มตัวอย่างที่ 2 ให้ตอบตามรูปแบบวิธีการตรวจของคู่มือ กลุ่มตัวอย่างที่ 3 ให้ตอบตามรูปแบบการตอบ 0 - 1 สำหรับการให้คะแนนการตอบแบบใหม่ให้นักศึกษาทำเครื่องหมาย + ลงในตัวเลือกที่เห็นว่าถูก และกาเครื่องหมาย - ลงในตัวเลือกที่เห็นว่าผิด และเว้นไว้ถ้าไม่แน่ใจ คะแนนของตัวเลือกแต่ละตัวจะได้จากคำตอบของผู้สอบ (+1, -1 หรือ 0) คะแนนที่ได้จากแต่ละข้อได้จากผลบวกของคะแนนในข้อต่าง ๆ ที่เห็นว่าผิด ถ้าทำเครื่องหมายผิดได้ถูกต้องจะได้ตัวเลือกละ 1 คะแนน แต่ถ้าทำเครื่องหมายบนตัวเลือกที่ถูกจะถูกหัก K - 1 คะแนน (K เท่ากับจำนวนตัวเลือกทั้งหมด) สำหรับวิธีการตอบแบบ 0 - 1 ให้ผู้สอบทำเครื่องหมายเฉพาะตัวเลือกที่ถูก ถ้าทำเครื่องหมายตัวเลือกที่ถูกได้ถูกต้องทั้งหมดได้ 1 คะแนนในแต่ละข้อ ซึ่งพบว่าคะแนนการเดาคะแนนข้อถูก คะแนนข้อผิดและคะแนนข้อที่ถูกเพียงบางส่วนของวิธีการตรวจให้คะแนนรูปแบบใหม่ ไม่แตกต่างจากวิธี 0 - 1 อย่างมีนัยสำคัญ และค่าความเชื่อมั่นวิธีใหม่สูงกว่าวิธีของคู่มือ และวิธี 0 - 1 อย่างมีนัยสำคัญ ส่วนค่าความเที่ยงตรงทั้ง 3 วิธีไม่แตกต่างกัน โดยทั่วไปวิธีการตอบ และตรวจให้คะแนนแบบใหม่มีประสิทธิภาพในการวัดเหนือกว่าวิธี 0 - 1 ซึ่งวิธีแบบใหม่ช่วยให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้น และลดจำนวนการเดาให้น้อยลงทำให้มีประสิทธิภาพในการวัดสูงและมีคุณค่าแก่การศึกษา

เพลินพิศ คุณคำ (2532 : 1 - 79) ศึกษาเปรียบเทียบค่าความเชื่อมั่นและปริมาณการเดาของแบบทดสอบปรนัยชนิดเลือกตอบที่มีการตอบและตรวจให้คะแนนวิธี 0 - 1 วิธีของคู่มือ และวิธีของอนันต์ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนปรังค์ภู จังหวัดศรีสะเกษ ภาคเรียนที่ 2 ปีการศึกษา 2531 จำนวน 30 คน ซึ่งเลือกมาโดยการสุ่มแบบธรรมดา (Simple Random Sampling) เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ (ว.204) บทที่ 10 เรื่องการใช้พลังงาน จำนวน 2 ฉบับ ฉบับที่ 1 มีจำนวน 20 ข้อ ฉบับที่ 2 มีจำนวน 16 ข้อ มีการตรวจให้คะแนนโดยวิธี 0 - 1, วิธีของคู่มือ และวิธีของอนันต์ ผลการศึกษาพบว่าการตรวจให้คะแนนวิธีของอนันต์ให้ค่าความเชื่อมั่นสูงกว่าวิธีอื่น ๆ ส่วนค่าความเที่ยงตรงของแบบทดสอบที่ตรวจให้คะแนนทั้ง 3 วิธี พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญ วิธีของอนันต์มีการเดาต่ำกว่าวิธี 0 - 1 และวิธีของคู่มือ และวิธีของอนันต์มีเปอร์เซ็นต์คะแนนที่ถูกกว่าวิธีของคู่มือ

จินตนา ต่อทรัพย์สินชัย (2533 : 46 - 49) ศึกษาค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยใช้กฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกัน 5 จุดกับการไม่กำหนดจุดซึลลิ่ง กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนสังกัดกรมสามัญศึกษา ในจังหวัดปทุมธานี ปีการศึกษา 2532 เครื่องมือที่ใช้ในการศึกษา คือ แบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ (ค. 311) ที่มีการจัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดจำนวน 50 ข้อ ผลการศึกษาพบว่า การตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 1, 2, 3 และ 4 ให้ค่าความเชื่อมั่นต่ำกว่าการตรวจที่กำหนดจุดซึลลิ่งจุดที่ 5 และการไม่กำหนดจุดซึลลิ่ง อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนการตรวจให้คะแนนโดยกำหนดจุดซึลลิ่งที่ 5 ให้ค่าความเชื่อมั่นสูงสุดเมื่อเทียบกับวิธีการตรวจแบบไม่กำหนดจุดซึลลิ่ง สำหรับค่าความเที่ยงตรงพบว่าการตรวจแบบไม่กำหนดจุดซึลลิ่งให้ค่าความเที่ยงตรงสูงสุด

อัมพร ยิ้มแย้ม (2533 : 59 - 63) ศึกษาผลการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกันที่มีต่อค่าการเดา และลักษณะโค้งอินฟอร์เมชันของแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดเพชรบุรี ปีการศึกษา 2533 เครื่องมือที่ใช้คือแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 เรื่องเลขยกกำลัง ที่มีการจัดเรียงข้อคำถามจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดจำนวน 40 ข้อ ผลการศึกษาพบว่า

1. ค่าการเดาของข้อสอบที่ได้จากการตรวจให้คะแนนที่จุดซึลลิ่งจุดที่ 1 ประมาณค่าไม่ได้ ค่าการเดาของข้อสอบที่ได้จากการตรวจให้คะแนนที่จุดซึลลิ่งจุดที่ 2 มีค่าต่ำสุด และค่าการเดาของข้อสอบที่ได้จากการตรวจให้คะแนนที่ไม่กำหนดจุดซึลลิ่งมีค่าสูงสุด ผลการเปรียบเทียบค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนในแต่ละจุดซึลลิ่งพบว่า ค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งที่ 2, 3, 4, และ 5 แต่ละจุดแตกต่างกับค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่ไม่กำหนดจุดซึลลิ่งอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าการเดาของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งระหว่างจุดที่ 2, 3, 4 และ 5 แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. จากการศึกษาโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนในแต่ละจุดซึลลิ่งพบว่าโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 2, 3, 4 และ 5 แต่ละจุดสูงกว่าโค้งอินฟอร์เมชันของแบบทดสอบ ที่ตรวจให้คะแนนที่ไม่กำหนดจุดซึลลิ่งในช่วงระดับความสามารถต่ำ โดยโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 2 สูงสุด ซึ่งหมายความว่าคะแนนที่ได้จากการตรวจที่กำหนดจุดซึลลิ่งจุดที่ 2 ใช้ประมาณค่าความสามารถได้ดีที่สุดเมื่อผู้สอบมีระดับความสามารถต่ำ และโค้งอินฟอร์เมชันของแบบทดสอบที่ตรวจให้คะแนนที่กำหนดจุดซึลลิ่งจุดที่ 5 สูงสุดในช่วงระดับความสามารถสูง ซึ่ง

หมายความว่าคะแนนที่ได้จากการตรวจที่กำหนดจุดซึลลิ่งจุดที่ 5 ใช้ประมาณค่าความสามารถได้ดีที่สุดเมื่อผู้สอบมีระดับความสามารถสูง

พินิจ อุไรรักษ์ (2533 : 57) ศึกษาผลของวิธีการให้คะแนนที่มีต่อคะแนนสอบต่อค่าความเชื่อมั่น และค่าความเที่ยงตรงจากการให้น้ำหนักคะแนนรายข้อเท่ากัน วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถาม และ วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในเขตจังหวัดบุรีรัมย์ ปีการศึกษา 2532 โดยสุ่มแบบหลายชั้นจำนวน 1,011 คน ผลการศึกษาพบว่าค่าความเชื่อมั่นของแบบสอบจากการให้น้ำหนักคะแนนรายข้อเท่ากัน วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถาม และวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบมีค่าเท่ากับ .705, .781, .886 และ .894 ตามลำดับ โดยเฉลี่ยวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามระดับความมั่นใจในการตอบ วิธีให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อคำถามและวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของผู้ตอบ มีค่าความเที่ยงตรงสูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน

เยาวลักษณ์ น้อยนาแสง (2535 : บทคัดย่อ) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่น ค่าความเที่ยงตรง อำนาจจำแนก เปอร์เซนต์การเดา เปอร์เซนต์ความถนัดทั้งหมด เปอร์เซนต์ความถนัดบางส่วน และเปอร์เซนต์ไม่มีความถนัดที่ใช้ในการตอบ และตรวจให้คะแนน 3 วิธี คือ วิธี 0 - 1 วิธีของคัมป์ส วิธีของอนันต์ กับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดร้อยเอ็ด ปีการศึกษา 2534 ผลการศึกษาพบว่าแบบทดสอบที่มีตัวเลือกถูกตัวเดียวมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มีตัวเลือกถูกหลายตัว เมื่อตรวจด้วยวิธี 0 - 1 และวิธีของคัมป์สมีความเที่ยงตรงสูงกว่าแบบ 0 - 1 ส่วนแบบทดสอบที่มีตัวเลือกถูกหลายตัว เมื่อตรวจด้วยวิธีของคัมป์ส และ วิธีอนันต์มีค่าความเชื่อมั่นสูงกว่าวิธี 0 - 1 ส่วนค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียน ที่มีตัวเลือกถูกตัวเดียว และมีตัวเลือกถูกหลายตัวเมื่อตรวจให้คะแนนทั้ง 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติโอกาสของการเดา เพศชายเดาน้อยกว่าเพศหญิง และ แบบทดสอบความถนัดทางการเรียนแบบที่มีตัวเลือกถูกตัวเดียวตรวจด้วยวิธีของอนันต์เปอร์เซนต์ความถนัดบางส่วนสูงกว่าวิธีของคัมป์ส แต่เปอร์เซนต์ความไม่ถนัดตรวจด้วยวิธีของอนันต์ต่ำกว่าวิธีของคัมป์ส

มานะ คิวเจริญ (2539 : 57 - 62) ศึกษาผลของการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกันของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยากและพฤติกรรมการเรียนรู้ โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดกาญจนบุรี ปีการศึกษา 2538 จำนวน 434 คน เครื่องมือที่ใช้เป็นแบบ

ทดสอบคณิตศาสตร์ (ค.204) ที่มีการจัดเรียงข้อคำถามต่างกันจำนวน 2 ฉบับ คือ จัดเรียงตามค่าความยากจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุดจำนวน 40 ข้อ และจัดเรียงตามพฤติกรรมการเรียนรู้จากชั้นความจำไปจนถึงชั้นการวิเคราะห์จำนวน 40 ข้อ เรื่อง สมการและอสมการ ร้อยละ และอัตราส่วน ปริมาตรและพื้นที่ผิว ผลการศึกษาพบว่า

1. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยากที่ได้จากการตรวจให้คะแนนโดยกฎซิลลิงที่กำหนดจุดซิลลิงตำแหน่งที่ 1 มีค่าความเชื่อมั่นสูงสุดและไม่กำหนดตำแหน่งจุดซิลลิงมีค่าความเชื่อมั่นต่ำสุด
2. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงตามพฤติกรรมการเรียนรู้ที่ได้จากการตรวจให้คะแนนโดยกฎซิลลิงที่กำหนดจุดซิลลิงตำแหน่งที่ 1 มีค่าความเชื่อมั่นสูงสุดและไม่กำหนดตำแหน่งจุดซิลลิงมีค่าความเชื่อมั่นต่ำสุด
3. ค่าความสอดคล้องระหว่างคะแนนของแบบทดสอบคณิตศาสตร์ที่มีการจัดเรียงตามค่าความยากกับเรียงตามพฤติกรรมการเรียนรู้ที่ตรวจให้คะแนนโดยกฎซิลลิง ที่กำหนดจุดซิลลิงตำแหน่งเดียวกันทุกตำแหน่ง มีค่าความสอดคล้องกันอย่างน้อยมีนัยสำคัญทางสถิติที่ระดับ .01 โดยแบบทดสอบที่ไม่กำหนดตำแหน่งจุดซิลลิงมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสูงสุด และจุดซิลลิงตำแหน่งที่หนึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนต่ำสุด

จากการศึกษางานวิจัยทั้งในประเทศและต่างประเทศ จะเห็นได้ว่า วิธีการตรวจให้คะแนนและวิธีการตอบมีผลต่อค่าความเชื่อมั่นของแบบทดสอบ ซึ่งการหาความเชื่อมั่นของแบบทดสอบส่วนมาก ไม่ว่าจะใช้วิธีการตรวจให้คะแนนและวิธีการตอบใด ๆ จะเป็นการหาความเชื่อมั่นทั้งฉบับ และวิธีการตรวจให้คะแนนแบบ 0 - 1 จะให้ค่าความเชื่อมั่นต่ำกว่าวิธีการตรวจให้คะแนนแบบความมั่นใจหรือให้น้ำหนักตามรายข้อ ส่วนวิธีการตรวจให้คะแนนโดยกฎซิลลิงแต่ละแบบนั้น จะหาความเชื่อมั่นของแบบทดสอบในแต่ละตำแหน่งของการตอบผิด ซึ่งในแต่ละตำแหน่งของการตอบผิดนั้น จะให้ค่าความเชื่อมั่นสูงกว่าค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ

7. งานวิจัยวิจัยที่เกี่ยวข้องกับการจัดเรียงข้อคำถาม

งานวิจัยในต่างประเทศ

มีผู้ศึกษาวิจัยเกี่ยวกับการจัดเรียงข้อคำถามไว้ ดังนี้

โมเลนคอฟ (Mollenkopf. 1950 : 291) ศึกษาค้นคว้าผลการจัดเรียงข้อคำถามในแบบทดสอบใหม่ว่าจะมีผลทำให้ค่าทางสถิติเปลี่ยนแปลงหรือไม่ และ เปรียบเทียบผลที่เกิดจากการสอบแบบไม่จำกัดเวลา (Power Test) กับการสอบแบบจำกัดเวลา (Speed Test) พบว่าค่าความยากง่ายของข้อสอบ และค่าอำนาจจำแนกของข้อสอบแต่ละข้อเปลี่ยนแปลงไปจากเดิม แสดงว่าการจัดเรียงข้อคำถามในแบบทดสอบมีผลทำให้ค่าทางสถิติของแบบทดสอบเปลี่ยนแปลงไป

แมคนิคอล (Flangher, Melton and Myers. 1968 : 813 ; citing Mac Nical. 1968) ศึกษาการเปลี่ยนวิธีในการเรียงลำดับข้อคำถามซึ่งแต่เดิมเรียงจากง่ายไปหายาก เปลี่ยนไปเป็นเรียงแบบสุ่ม ผลการศึกษาพบว่าการเรียงลำดับข้อคำถามจากข้อยากไปหาข้อง่ายทำให้ข้อสอบยากขึ้นกว่าเดิมอย่างมีนัยสำคัญทางสถิติ แต่การเรียงแบบสุ่มไม่มีผลทำให้ค่าความยากง่ายของข้อสอบแตกต่างไปจากเดิม

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

1. จัดเรียงข้อสอบแบบมาตรฐานเดิม (Standard Arrangement)
2. จัดเรียงข้อภายในแต่ละตอน ซึ่งมีตอนละ 5 ข้อใหม่ ส่วนการเรียงระหว่างตอนยังเหมือนมาตรฐานเดิม
3. จัดเรียงลำดับข้อสอบแต่ละตอนใหม่ ส่วนข้อภายในแต่ละตอนนั้นเรียงเหมือนมาตรฐานเดิม
4. จัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอน

ผลการวิจัยพบว่าการจัดเรียงแบบมาตรฐานเดิม และการจัดเรียงข้อภายในแต่ละตอนแบบมาตรฐานเดิมไม่มีผลทำให้คะแนนในการสอบเปลี่ยนแปลง ส่วนการจัดเรียงลำดับข้อสอบแต่ละตอนเสียใหม่ข้อภายในแต่ละตอนนั้นเรียงเหมือนมาตรฐานเดิม และการจัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอนมีผลทำให้คะแนนในการสอบเปลี่ยนแปลงไป

คลอสเนอร์ และแกลล์แมน (Klosner and Gellman 1973 : 413 - 418) ได้ศึกษาการเรียงลำดับข้อคำถามในแบบทดสอบ โดยเรียงตามหัวเรื่องตามลำดับจากง่ายไปยาก และเรียงตามหัวเรื่องซึ่งแต่ละหัวเรื่องเรียงข้อคำถามจากง่ายไปหายาก โดยทดลองกับกลุ่มตัวอย่างที่เป็นนักศึกษาของมหาวิทยาลัยฮอฟสตรา ปี ค.ศ. 1969 จำนวน 54 คน แบ่งออกเป็น 2 กลุ่ม คือ กลุ่มที่มีความสามารถสูง กับกลุ่มที่มีความสามารถต่ำ เครื่องมือที่ใช้เป็นแบบทดสอบวิชาวัดผลการศึกษาเบื้องต้นจำนวน 75 ข้อ ผลการศึกษาพบว่าในกลุ่มนักศึกษาที่มีความสามารถต่ำการเรียงลำดับข้อคำถามของแบบทดสอบมีผลทำให้คะแนนเปลี่ยนไป ส่วนค่าความเชื่อมั่นของแบบทดสอบที่เรียงตามหัวเรื่อง และเรียงตามลำดับจากข้อง่ายไปหาข้อยากแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เพลค (Plake. 1980 : 56 - 58) ศึกษาเกี่ยวกับการจัดเรียงข้อคำถาม 3 แบบ คือ เรียงจากง่ายไปหายาก เรียงแบบยากง่ายหมุนเวียน และ เรียงแบบสุ่ม โดยศึกษากับกลุ่มตัวอย่างที่

เป็นนักเรียนพยาบาลจำนวน 104 คน พบว่าคะแนนการจัดเรียงข้อคำถามทั้ง 3 แบบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

เลนี เจอาร์ และบูล (Lane, JR & Bull. 1987 : 865 - 878) ศึกษาผลของความรู้ เพศ ความยากเชิงสถิติ และความยากเชิงปัญญา ที่มีต่อการทำข้อสอบที่เรียงข้อคำถามแตกต่างกัน ตามความยากเชิงสถิติ และเรียงตามความยากเชิงปัญญา คือ จำแนกตามพฤติกรรมการเรียนรู้ ของบลูมชั้นความรู้ความจำ ความเข้าใจ และการนำไปใช้ เครื่องมือที่ใช้เป็นแบบทดสอบ แบบเลือกตอบ โดยศึกษากับกลุ่มตัวอย่างที่เป็นนักศึกษา 2 กลุ่ม กลุ่มแรกมีจำนวน 155 คน กลุ่มที่สองมีจำนวน 247 คน ทำการศึกษา 2 ครั้ง ครั้งที่ 1 ใช้แบบทดสอบ 5 ฉบับที่เรียง ข้อคำถามต่างกัน ผลการศึกษาพบว่าคะแนนรวมที่ได้จากแบบทดสอบของเพศชายกับเพศหญิง แตกต่างกัน และข้อสอบแต่ละฉบับจัดทำเป็น 2 ชุด ชุดแรกจะมีข้อคำถามบอกระดับความยาก เชิงปัญญาอยู่ทางด้านซ้ายมือของข้อสอบแต่ละข้อ เพื่อให้ผู้สอบใช้เป็นประโยชน์ในการทำ ข้อสอบ ส่วนข้อสอบอีกชุดหนึ่งจะไม่มีข้อความบอกระดับความยากเชิงปัญญา ผลการศึกษา พบว่าข้อสอบชุดแรกไม่มีผลทำให้คะแนนรวมของเพศชายและเพศหญิง แตกต่างกันอย่างมี นัยสำคัญทางสถิติ แต่คะแนนที่ได้จากแบบทดสอบที่มีการเขียนข้อความบอกระดับความยาก เชิงปัญญากับคะแนนที่ได้จากแบบทดสอบที่ไม่มีข้อความบอกระดับความยากเชิงปัญญา แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

งานวิจัยภายในประเทศ

เดือนใจ เศรษฐลักโก (2511 : 45 - 48) ศึกษาการจัดระเบียบข้อทดสอบวิชาภาษาอังกฤษ ภาษาไทย วิทยาศาสตร์ และ คณิตศาสตร์ โดยจัดเรียงข้อสอบด้วยวิธีที่แตกต่างกัน 4 วิธี ในแต่ละวิธีทำเป็น 2 ชุด ใช้เวลาในการสอบแตกต่างกัน รวมมีแบบทดสอบทั้งหมด 8 ชุด คือ ชุดที่ 1 เรียงแบบทดสอบจากฉบับวิชาภาษาอังกฤษ ภาษาไทย วิทยาศาสตร์ และ คณิตศาสตร์ ตามลำดับ ในแต่ละฉบับเรียงข้อสอบจากง่ายไปยาก ชุดที่ 2 เรียงแบบทดสอบเหมือนชุดที่ 1 แต่ เพิ่มเวลาสอบอีกฉบับละ 10 นาที ชุดที่ 3 เรียงแบบทดสอบจากวิชาคณิตศาสตร์ วิทยาศาสตร์ ภาษาไทย และ ภาษาอังกฤษ ในแต่ละฉบับเรียงข้อสอบจากง่ายไปยาก ชุดที่ 4 เรียงเหมือนชุด ที่ 3 แต่เพิ่มเวลาในการสอบฉบับละ 10 นาที ชุดที่ 5 เรียงแบบทดสอบจากฉบับภาษาอังกฤษ ภาษาไทย วิทยาศาสตร์และคณิตศาสตร์ในแต่ละฉบับเรียงข้อสอบแบบสลับ ชุดที่ 6 เรียงเหมือน ชุดที่ 5 แต่เพิ่มเวลาในการสอบอีกฉบับละ 10 นาที ชุดที่ 7 เรียงแบบผสมจัดทำเป็น 5 ฉบับ ย่อยแต่ละฉบับย่อยจะมีข้อสอบทุกวิชา ชุดที่ 8 เรียงเหมือนชุดที่ 7 แต่เพิ่มเวลาในการสอบอีก ฉบับละ 9 นาที โดยศึกษากับนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนดาราคาม จำนวน 240 คน แบ่งเป็น 8 กลุ่ม ๆ ละ 30 คน โดยวิธีแมชกรุป (Matched Group) ผลการวิจัยพบว่า การเรียงข้อสอบต่างกัน มีผลทำให้ค่าความเชื่อมั่นของข้อทดสอบวิชาภาษาอังกฤษแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยค่าความเชื่อมั่นที่เรียงข้อทดสอบแบบผสมมีค่าต่ำสุด

ส่วนการเรียงข้อทดสอบที่แตกต่างกัน มีผลทำให้ค่าความเชื่อมั่นของข้อทดสอบวิชาคณิตศาสตร์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยความเชื่อมั่นของข้อทดสอบที่เรียงแบบ ผสมมีค่าต่ำสุด

นิโลบล นิมกังรัตน์ (2519 : 56 - 58) ได้ศึกษาอิทธิพลของการจัดเรียงลำดับข้อคำถาม ที่จำแนกตามสมรรถภาพสมองที่มีต่อคุณภาพของแบบทดสอบคณิตศาสตร์ และภาษาไทยใน ระดับชั้นมัธยมศึกษาปีที่ 3 กับทุกโรงเรียนในเขตเทศบาลจังหวัดนครสวรรค์ ปีการศึกษา 2518 ผลการศึกษาพบว่า

1. การจัดเรียงลำดับข้อคำถามที่จำแนกตามสมรรถภาพสมองที่จะวัดของแบบทดสอบ ภาษาไทยทั้ง 6 วิธีมีผลต่อคะแนนการสอบของนักเรียน และคุณภาพของแบบทดสอบด้านความ เชื่อมั่น ความเที่ยงตรง และความยากมาตรฐานแตกต่างกัน แต่ไม่มีผลทำให้ค่าอำนาจจำแนก แตกต่างกันอย่างมีนัยสำคัญ

2. การจัดเรียงลำดับข้อคำถามที่จำแนกตามสมรรถภาพสมองที่จะวัดของแบบทดสอบ คณิตศาสตร์ทั้ง 6 วิธี ไม่มีผลทำให้คะแนนการสอบของนักเรียนและคุณภาพของแบบทดสอบ ทุกด้านแตกต่างกันอย่างมีนัยสำคัญ

3. จากการศึกษาพบว่าแบบทดสอบภาษาไทยชุดที่จัดเรียงข้อคำถามที่วัดสมรรถภาพ สมองด้านความรู้ความจำให้คะแนนการสอบของนักเรียน และ คุณภาพของแบบทดสอบที่ดีที่สุด ส่วนชุดที่เรียงข้อสอบวัดสมรรถภาพสมองด้านการนำไปใช้ จะให้คะแนนการสอบของนักเรียน และคุณภาพของแบบทดสอบต่ำสุด

4. แนวโน้มของการเปลี่ยนแปลงของคุณภาพและคะแนนการสอบของนักเรียนจากแบบ ทดสอบคณิตศาสตร์ และภาษาไทย มีลักษณะเกือบตรงกันข้าม ยกเว้นด้านความยากมาตรฐาน ที่มีลักษณะคล้ายคลึงกัน

5. แบบทดสอบภาษาไทยที่ควรใช้กับกลุ่มที่มีความถนัดทางการเรียนสูงและต่ำเมื่อได้ พิจารณาถึงความเหมาะสมของแบบทดสอบทั้งในด้านคุณภาพของแบบทดสอบคะแนนการสอบ ของนักเรียนและลักษณะของการแจกแจงที่เป็นไปตามทฤษฎีแล้วพบว่า ควรใช้แบบทดสอบที่ มีการจัดเรียงข้อคำถาม ดังนี้ คือ วัดความรู้ความจำ - ความเข้าใจ - การนำไปใช้หรือสูงกว่า มีแนวโน้มจะดีและเหมาะสมที่สุด

6. แบบทดสอบภาษาไทยและคณิตศาสตร์ ที่ควรใช้กับกลุ่มนักเรียนที่มีความถนัดทาง การเรียนปานกลางพบว่าใช้แบบทดสอบที่มีการเรียงแบบใดก็ได้

7. แบบทดสอบคณิตศาสตร์ที่ควรใช้กับกลุ่มที่มีความถนัดทางการเรียนสูง และต่ำเมื่อ ได้พิจารณาถึงความเหมาะสมกับแบบทดสอบทั้งในด้านคุณภาพของแบบทดสอบ คะแนนการ สอบของนักเรียนและลักษณะโค้งของการแจกแจงที่เป็นไปตามทฤษฎี พบว่าควรใช้แบบทดสอบ

ที่มีการจัดเรียงข้อคำถาม ดังนี้ คือ วัดความเข้าใจ - การนำไปใช้หรือสูงกว่า - ความรู้ความจำ มีแนวโน้มจะดีและเหมาะสมที่สุด

ศศิธร สุวรรณสุข (2519 : 72 - 90) ศึกษาการจัดเรียงลำดับข้อคำถามที่ต่างกันจะส่งผลกระทบต่อความเที่ยงตรง และค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 ปีการศึกษา 2518 สังกัดองค์การบริหารส่วนจังหวัดสมุทรปราการ จำนวน 200 คน แบ่งเป็นนักเรียนที่มีความสามารถทางคณิตศาสตร์สูง และนักเรียนที่มีความสามารถทางคณิตศาสตร์ต่ำ แล้วสุ่มนักเรียนในแต่ละระดับเข้ากลุ่ม ๆ ละ 25 คน เครื่องมือที่ใช้เป็นแบบทดสอบวิชาคณิตศาสตร์ 3 ฉบับ คือ ทักษะทางคณิตศาสตร์ (40 ข้อ) การแก้ปัญหาทางคณิตศาสตร์ (30 ข้อ) การหาเหตุผลทางคณิตศาสตร์ (30ข้อ) ซึ่งแบบทดสอบแต่ละฉบับมีการจัดเรียงลำดับข้อคำถาม 4 แบบ ได้แก่ การเรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อยากไปหาข้อง่าย เรียงตามหัวข้อเนื้อหาวิชา และเรียงจากข้อง่ายไปหาข้อยากในแต่ละหัวข้อเนื้อหาวิชา ผลจากการศึกษาพบว่า

1. เมื่อพิจารณานักเรียนทั้งหมด ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะทางคณิตศาสตร์ที่มีการเรียงข้อสอบทั้ง 4 แบบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. เมื่อพิจารณานักเรียนที่มีความสามารถทางคณิตศาสตร์สูง ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะทางคณิตศาสตร์ที่มีการเรียงข้อสอบทั้ง 4 แบบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
3. เมื่อพิจารณานักเรียนที่มีความสามารถทางคณิตศาสตร์ต่ำค่าความเชื่อมั่นของแบบทดสอบวัดทักษะทางคณิตศาสตร์ การแก้ปัญหา และการหาเหตุผลทางคณิตศาสตร์ที่มีการเรียงข้อสอบทั้ง 4 แบบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05
4. ค่าความเชื่อมั่นของแบบทดสอบวัดการแก้ปัญหาทางคณิตศาสตร์ ที่เรียงตามหัวข้อเนื้อหาวิชาของนักเรียนกลุ่มความสามารถทางคณิตศาสตร์สูง และนักเรียนกลุ่มความสามารถทางคณิตศาสตร์ต่ำแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สิริพร พุ่มแก้ว (2521 : 30 - 39) ศึกษาผลการเรียงลำดับข้อสอบต่อค่าความเชื่อมั่นค่าความเที่ยงตรงของแบบสอบ และผลสัมฤทธิ์ทางการเรียน โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน เครื่องมือที่ใช้แบบทดสอบผลสัมฤทธิ์วิชาภาษาไทย และ วิชาคณิตศาสตร์ที่มีการจัดเรียงข้อสอบ 4 แบบ คือ เรียงจากง่ายไปยาก เรียงจากยากมาง่าย เรียงตามเนื้อหา และ เรียงแบบสุ่ม ผลการศึกษาพบว่าการจัดเรียงข้อสอบทั้ง 4 แบบมีผลให้ค่าความเชื่อมั่นของแบบสอบเปลี่ยนแปลงน้อยมาก

แนนน้อย เพชรรัตน์ (2521 : 36 - 40) ศึกษาว่าการเรียงลำดับข้อคำถามของแบบทดสอบด้วยวิธีต่างกัน คือ เรียงจากง่ายไปยาก เรียงตามลำดับความยากแบบหมุนเวียน และแบบสุ่ม โดยมีได้คำนึงถึงเนื้อหาของข้อคำถามนั้นว่าจะมีผลต่อคุณภาพของแบบทดสอบวิชา

วิทยาศาสตร์ และแบบทดสอบทักษะขบวนการทางวิทยาศาสตร์อย่างไร กลุ่มตัวอย่างที่ศึกษา เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2520 ในเขตจังหวัดนครปฐม จำนวน 300 คน ผลจากการศึกษาพบว่า การเรียงลำดับข้อคำถามแบบสุ่มในแบบทดสอบทักษะขบวนการทางวิทยาศาสตร์ส่งผลต่อความเชื่อมั่นสูงกว่าการเรียงคำถามแบบอื่น ๆ ส่วนการเรียงลำดับข้อคำถามที่เรียงจากง่ายไปยาก เรียงตามลำดับความยากแบบหมุนเวียน ส่งผลต่อความเที่ยงตรง ความเชื่อมั่น ความยากง่าย และอำนาจจำแนกในแบบทดสอบทั้ง 2 ฉบับไม่แตกต่างกัน

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

ผลการศึกษาพบว่า

1. ความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ ที่เรียงข้อคำถามตามค่าความยาก จากค่าสูงไปค่าต่ำจะมีค่าความเชื่อมั่นสูงกว่า เรียงข้อคำถามตามค่าความยากจากค่าต่ำไปค่าสูง และเมื่อทดสอบความแตกต่างพบว่าต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ ที่เรียงข้อคำถามตามพฤติกรรม จากขั้นต่ำไปขั้นสูง มีค่าความเชื่อมั่นสูงกว่าที่เรียงข้อคำถามตามพฤติกรรมจากขั้นสูงไปขั้นต่ำ และเมื่อทดสอบความแตกต่างพบว่าต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ความเชื่อมั่นของแบบทดสอบคณิตศาสตร์ ที่เรียงตามพฤติกรรมจากขั้นต่ำไปขั้นสูง และเรียงข้อคำถามภายในแต่ละพฤติกรรมตามค่าความยากจากค่าต่ำไปค่าสูง มีค่าความเชื่อมั่น สูงที่สุด และเมื่อทดสอบความแตกต่างพบว่ามีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ศิริวิมล สุขสมบูรณ์ (2541 : 69) ศึกษาผลการวิเคราะห์แบบแผนการตอบข้อสอบของ แบบทดสอบเลือกตอบที่มีรูปแบบการจัดเรียงข้อคำถามต่างกัน คือเรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อง่ายไปหาข้อยากตามเนื้อหาบทเรียน เรียงตามพฤติกรรมที่ต้องการวัด เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือกวิชาเศรษฐศาสตร์เบื้องต้น จำนวน 40 ข้อเกี่ยวกับ ความรู้เบื้องต้นทางเศรษฐศาสตร์ ระบบเศรษฐกิจ การผลิต และการบริโภคกลุ่ม กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 2 ปีการศึกษา 2540 สถาบันเทคโนโลยี ราชมนคลในเขตกรุงเทพมหานครที่ทำการสอนในสาขาวิชาบริหารธุรกิจ จำนวน 5 % จาก 5 วิทยาเขตได้มาโดยการสุ่มแบบแบ่งชั้นจำนวน 1,276 คน ผลการศึกษาพบว่าค่าความเชื่อมั่น ของแบบทดสอบด้วยวิธีแบ่งครึ่งฉบับมีค่าตั้งแต่ 0.6270 ถึง 0.8253 โดยแบบทดสอบที่มีรูปแบบ การจัดเรียงข้อคำถามแบบเรียงจากข้อง่ายไปหาข้อยากมีค่าความเชื่อมั่นต่ำสุดและแบบทดสอบ ที่มีรูปแบบการจัดเรียงข้อคำถาม แบบเรียงจากข้อง่ายไปหาข้อยากตามเนื้อหาของบทเรียนมีค่า ความเชื่อมั่นสูงสุด ส่วนค่าความเชื่อมั่นของแบบทดสอบด้วยวิธีสอบซ้ำมีค่าตั้งแต่ 0.5530 ถึง

0.6856 โดยแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถามแบบเรียงจากข้อง่ายไปหาข้อยากมีค่าความเชื่อมั่นต่ำสุด และแบบทดสอบที่มีรูปแบบการจัดเรียงข้อคำถามตามพฤติกรรมที่ต้องการวัดมีค่าความเชื่อมั่นสูงสุด ส่วนค่าความเชื่อมั่นของแบบทดสอบทั้ง 3 ฉบับที่คำนวณด้วยวิธีแบ่งครึ่งฉบับและวิธีสอบซ้ำมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บังอร กมลวัฒนา (2542 : 56 - 58) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบ และการจัดเรียงปัญหาต่างกัน เครื่องมือที่ใช้เป็นแบบทดสอบวัดความสามารถในการแก้ปัญหา ตามกระบวนการแก้ปัญหาของเวียร์ (Weir) จำนวน 4 ฉบับ คือ 1) แบบทดสอบเลือกตอบที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยากชนิด 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 โดยศึกษากับนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 สังกัดคณะกรรมการการศึกษาเอกชน กรุงเทพมหานคร จำนวน 1,600 คน ผลการศึกษาพบว่าค่าความเชื่อมั่นแบบเลือกตอบที่เรียงข้อปัญหาตามความยากมาตรฐานเฉลี่ยจากง่ายไปยาก กับแบบเลือกตอบที่เรียงข้อปัญหาโดยวิธีสุ่มแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และค่าความเชื่อมั่นที่จัดเรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยากแบบเลือกตอบ กับการจัดเรียงข้อปัญหาตามความยากมาตรฐานเฉลี่ยจากง่ายไปยากแบบตอบสั้นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

รัชณี กันภัย (2543 : 77 - 78) ศึกษาค่าความเชื่อมั่น และค่าความเที่ยงตรงของแบบทดสอบวัดความถนัดด้านคณิตศาสตร์ที่สอบด้วยไมโครคอมพิวเตอร์ จำนวน 4 แบบ ๆ ละ 50 ข้อโดยแบบทดสอบแต่ละแบบจะมีคำถามเหมือนกันทุกข้อ ค่าความเชื่อมั่นของแบบทดสอบที่มีการจัดเรียงข้อสอบ และ การกำหนดเวลาให้ทำข้อสอบวิเคราะห์โดยใช้สัมประสิทธิ์ r_B เท่ากับ 0.88, 0.92, 0.81 และ 0.89 ตามลำดับ

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

1. ค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบตามค่าความยากโดยกำหนดเวลาให้ทำข้อสอบเป็นรายข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบตามค่าความยากโดยกำหนดเวลาให้ทำข้อสอบรวมทั้งฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ค่าความเชื่อมั่นของแบบทดสอบ ที่จัดเรียงข้อสอบแบบสุ่มโดยกำหนดเวลาให้ทำข้อสอบเป็นรายข้อกับค่าความเชื่อมั่นของแบบทดสอบที่จัดเรียงข้อสอบแบบสุ่มโดยกำหนดเวลาให้ทำข้อสอบรวมทั้งฉบับมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำเป็นรายข้อเมื่อจัดเรียงข้อสอบตามค่าความยากกับค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำเป็นรายข้อเมื่อจัดเรียงข้อสอบแบบสุ่มมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4. ค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำรวมทั้งฉบับ เมื่อจัดเรียงข้อสอบตามค่าความยาก กับค่าความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำรวมทั้งฉบับเมื่อจัดเรียงข้อสอบแบบสุ่มมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากเอกสารงานวิจัยทั้งในประเทศและต่างประเทศแสดงให้เห็นว่า การจัดเรียงข้อคำถามของแบบทดสอบมีผลต่อค่าความเชื่อมั่นคือ แบบทดสอบจะมีค่าความเชื่อมั่นสูง เมื่อเรียงข้อคำถามจากง่ายไปยาก

จะเห็นได้ว่า ทั้งวิธีการตรวจให้คะแนนการจัดเรียงข้อคำถามมีผลต่อค่าความเชื่อมั่นของแบบทดสอบ ผู้วิจัยจึงมีความสนใจที่จะศึกษาวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบ จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีการเรียงข้อคำถามตามค่าความยากจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุดว่าจะมีค่าความเชื่อมั่นแตกต่างกันหรือไม่และการตอบผิดตำแหน่งใดมีผลต่อค่าความเชื่อมั่นของแบบทดสอบสูงสุด เนื่องจากยังไม่มีผู้ใดเคยศึกษาเรื่องนี้ในระดับประถมศึกษาศึกษามาก่อน

8. สมมติฐานในการศึกษาค้นคว้า

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน

2. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน

3. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน

4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกัน

5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สองมีค่าแตกต่างกัน

6. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบ
ของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษา

ในการศึกษาครั้งนี้ ผู้วิจัยต้องการศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่เรียงข้อคำถามตามค่าความยากจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุดที่ตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกัน 3 แบบ ซึ่งมีวิธีดำเนินการศึกษาตามรายละเอียด ดังนี้

ประชากร

ประชากรที่ใช้ในการศึกษา ได้แก่นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ จำนวน 444 โรงเรียน 490 ห้องเรียน และมีนักเรียนจำนวน 9,684 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษา ได้แก่นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ เลือกมาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) สุ่มได้ 35 ห้องเรียน นักเรียนจำนวน 790 คน ตามขั้นตอนการสุ่ม ดังนี้

1. ประมาณขนาดกลุ่มตัวอย่างที่ใช้ในการวิจัยด้วยความเชื่อมั่น 95 เปอร์เซนต์ ($\alpha = .05$) เมื่อเทียบจากตารางขนาดกลุ่มตัวอย่างต่ำสุดที่ควรเลือกจากประชากรเท่ากับ 385 คน (ล้วน สายยศ และอังคณา สายยศ. 2540 : 400 ; อ้างอิงมาจาก Yamane. 1967 : 886 - 887)

2. แบ่งโรงเรียนออกเป็น 3 ขนาด คือ โรงเรียนขนาดใหญ่ 6 โรงเรียน 19 ห้องเรียน โรงเรียนขนาดกลาง 88 โรงเรียน 123 ห้องเรียน และโรงเรียนขนาดเล็ก 350 โรงเรียน 348 ห้องเรียน

3. ใช้สัดส่วนร้อยละ 7 สุ่มโรงเรียนแต่ละขนาดในข้อที่ 2 มาโรงเรียนละ 1 ห้องเรียน ดังแสดงไว้ในตาราง 6

ตาราง 6 จำนวนนักเรียนกลุ่มตัวอย่างที่ใช้ในการทดสอบจำแนกตามขนาดโรงเรียน

ขนาดโรงเรียน	ชื่อโรงเรียน	จำนวนกลุ่มตัวอย่าง	
		ห้องเรียน	นักเรียน
โรงเรียนขนาดใหญ่	1. โรงเรียนวัดศรีโลหะราษฎร์บำรุง	1	41
	2. โรงเรียนอนุบาลกาญจนบุรี	1	46
โรงเรียนขนาดกลาง	1. โรงเรียนบ้านสระลุมพุก	1	19
	2. โรงเรียนวัดวังศาลา	1	25
	3. โรงเรียนวัดห้วยเหินยว	1	27
	4. โรงเรียนวัดหนองเสือ	1	27
	5. โรงเรียนอนุบาลพนมทวน	1	28
	6. โรงเรียนวัดท่าเรือ	1	30
	7. โรงเรียนวัดพระแท่นดงรัง	1	31
	8. โรงเรียนวัดดอนชะเอม	1	35
	9. โรงเรียนบ้านท่ามะกา	1	36
โรงเรียนขนาดเล็ก	1. โรงเรียนบ้านหนองดาดง	1	8
	2. โรงเรียนบ้านหนองตาแพ่ง	1	11
	3. โรงเรียนวัดหมอสอวังคา	1	12
	4. โรงเรียนอุโลกสีหมีน	1	12
	5. โรงเรียนวัดเขาตะพั่น	1	13
	6. โรงเรียนหนองหิน	1	13
	7. โรงเรียนวัดหนองพันท้าว	1	14
	8. โรงเรียนบ้านท่าพะเนียง	1	14
	9. โรงเรียนบ้านมะกอกหมู	1	17
	10. โรงเรียนวัดเขาใหญ่	1	17
	11. โรงเรียนวัดตาปานนิมิต	1	19
	12. โรงเรียนบ้านหนองซ้นฝิ่ง	1	19
	13. โรงเรียนบ้านไร่ร่วมวิทยาคาร	1	19
	14. โรงเรียนวัดหนองพลับ	1	20
	15. โรงเรียนบ้านหนองลาน	1	20

ตาราง 6 (ต่อ)

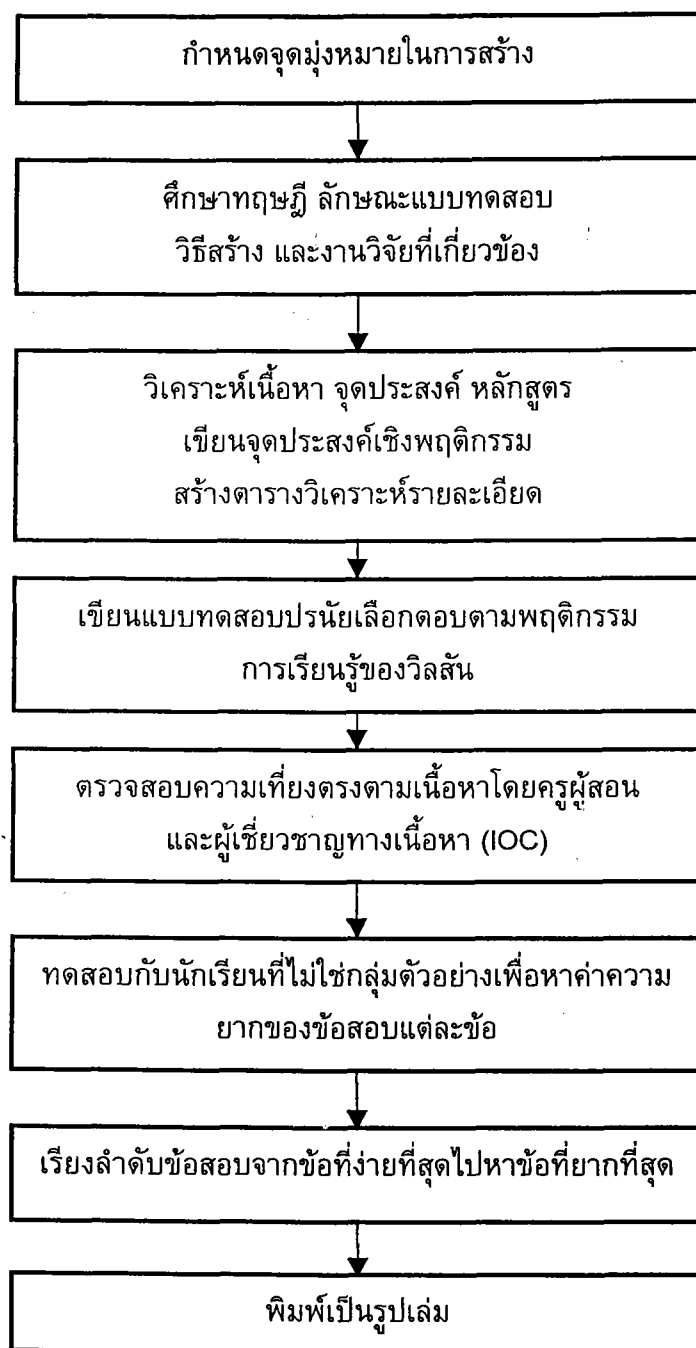
ขนาดโรงเรียน	ชื่อโรงเรียน	จำนวนกลุ่มตัวอย่าง	
		ห้องเรียน	นักเรียน
โรงเรียนขนาดเล็ก	16. โรงเรียนบ้านทุ่งประทุน	1	22
	17. โรงเรียนวัดทุ่งมะกรูด	1	22
	18. โรงเรียนบ้านดอนเขว้า	1	23
	19. โรงเรียนวัดดอนขมิ้น	1	23
	20. โรงเรียนวัดแสนตอ (วรวัตตวิทยาการ)	1	24
	21. โรงเรียนวัดใหม่เจริญผล	1	24
	22. โรงเรียนวัดดงสักหมื่นวิทยาการ	1	26
	23. โรงเรียนวัดหนองโรง	1	26
	24. โรงเรียนวัดหนองลาน	1	27
	รวม	35	790

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ตามพฤติกรรมการเรียนรู้ของวิลสัน (Wilson 1971 : 645 - 696) ซึ่งวัดความรู้ด้านการคิดคำนวณ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ เนื้อหาบทที่ 1 - 8 ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ซึ่งผู้วิจัยสร้างขึ้นเองจำนวน 50 ข้อ

วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้มีขั้นตอนการสร้าง คือ



ภาพประกอบขั้นตอนการดำเนินงาน

รายละเอียดขั้นตอนการสร้างเครื่องมือ

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบปรนัยแบบเลือกตอบวิชาคณิตศาสตร์ที่จัดเรียงข้อคำถามตามค่าความยาก

2. ศึกษาทฤษฎี เอกสาร ตำรา ลักษณะแบบทดสอบ วิธีสร้าง และงานวิจัยที่เกี่ยวข้อง

3. ศึกษาหลักสูตรวิชาคณิตศาสตร์ประถมศึกษาตอนปลาย พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) วิเคราะห์เนื้อหา จุดประสงค์การเรียนรู้ใน ป.02 และ คู่มือครูการสอนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) สร้างตารางวิเคราะห์รายละเอียดโดยใช้เนื้อหาบทที่ 1 - 8 ดังแสดงในตาราง 7

ตาราง 7 ตารางการวิเคราะห์รายละเอียดวิชาคณิตศาสตร์ เนื้อหา บทที่ 1 - 8
 ชั้นประถมศึกษาปีที่ 6

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม จำนวน ข้อ
	การคิด คำนวณ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
บทที่ 1 จำนวนและการบวก ลบ คูณ หาร					
1. เมื่อกำหนดตัวเลขแทนจำนวนนับให้นักเรียนสามารถบอกค่าของตัวเลขตามค่าประจำหลักได้	1	-	-	-	1
2. เมื่อกำหนดตัวเลขแทนจำนวนนับหลาย ๆ จำนวนให้นักเรียนสามารถเรียงลำดับจำนวนที่มีค่าน้อยไปหาจำนวนที่มีค่ามากหรือจากจำนวนที่มีค่ามากไปหาจำนวนที่มีค่าน้อยได้	-	1	-	-	1
3. เมื่อกำหนดจำนวนนับให้นักเรียนสามารถหาค่าประมาณใกล้เคียงจำนวนเต็มได้	-	1	1	-	2
4. เมื่อกำหนดโจทย์หรือโจทย์ปัญหาการบวก ลบ คูณ หาร ให้นักเรียนสามารถหาคำตอบได้	-	1	2	1	4
บทที่ 2 สมการและการแก้สมการ					
1. เมื่อกำหนดประโยคสัญลักษณ์ให้หลาย ๆ ประโยค นักเรียนสามารถบอกได้ว่าประโยคใดเป็นสมการ	1	-	-	-	1
2. เมื่อกำหนดสมการให้หลาย ๆ สมการนักเรียนสามารถบอกได้ว่าสมการใดเป็นสมการที่เป็นจริง	-	1	-	-	1
3. เมื่อกำหนดสมการให้นักเรียนสามารถทำให้สมการเป็นจริงได้	-	1	-	-	1

ตาราง 7 (ต่อ)

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม จำนวน ข้อ
	การคิด คำนวณ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
4. เมื่อกำหนดโจทย์ปัญหาในชีวิตประจำวันที่ต้องใช้การบวก ลบ คูณหรือหาร อย่างใดอย่างหนึ่งให้ นักเรียนสามารถแปลงโจทย์ปัญหาให้อยู่ในรูปสมการและหาคำตอบได้	-	1	3	1	5
บทที่ 3 ตัวประกอบของจำนวนนับ					
1. เมื่อกำหนดจำนวนนับให้นักเรียนสามารถหาตัวประกอบได้	-	1	1	-	2
2. เมื่อกำหนดจำนวนนับให้ นักเรียนสามารถหาจำนวนเฉพาะและตัวประกอบเฉพาะได้	1	1	-	-	2
3. เมื่อกำหนดจำนวนนับให้ นักเรียนสามารถแยกตัวประกอบได้	1	1	-	-	2
4. เมื่อกำหนดจำนวนนับที่ไม่เกิน 4 จำนวนให้นักเรียนสามารถหา ห.ร.ม.และ ค.ร.น.ได้	-	-	3	1	4
บทที่ 4 มุมและส่วนของเส้นตรง					
1. นักเรียนสามารถบอกชนิดของมุมได้	1	-	-	-	1
2. นักเรียนสามารถหามุมที่เท่ากับมุมที่กำหนดให้ได้	-	1	2	-	3
บทที่ 5 เส้นขนาน					
1. เมื่อกำหนดเส้นขนานที่มีเส้นตัด นักเรียนสามารถบอกชื่อมุมแย้งที่เท่ากันและชื่อมุมภายในที่อยู่ข้างเดียวกันของเส้นตัดที่รวมกันได้ 180 องศา	1	-	1	-	2
2. เมื่อกำหนดรูปให้นักเรียนสามารถใช้คุณสมบัติขนาดของมุมแย้งเท่ากัน และขนาดของมุมภายในที่อยู่บนข้างเดียวกันของเส้นตัดรวมกันได้ 180 องศา เพื่อบอกว่าเส้นตรงคู่ใดขนานกัน	1	1	-	1	3

ตาราง 7 (ต่อ)

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม จำนวน ข้อ
	การคิด คำนวณ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
บทที่ 6 ทิศและแผนผัง					
1. เมื่อกำหนดทิศทั้ง 8 ให้ นักเรียนสามารถบอกได้ว่า ทิศใดทำมุมกี่องศากับอีกทิศหนึ่ง	1	-	-	-	1
2. เมื่อกำหนดตำแหน่งสองตำแหน่งให้นักเรียนสามารถบอกได้ว่าตำแหน่งหนึ่งอยู่ทางทิศใดของอีกตำแหน่งหนึ่ง	-	1	-	-	1
3. เมื่อกำหนดแผนผังให้ นักเรียนสามารถบอกความยาวจริงหรือระยะทางจริงและรูปร่างหรือตำแหน่งของสิ่งต่างๆ ที่อยู่ในแผนผังได้	-	-	1	-	1
4. เมื่อกำหนดแผนผังแสดงการเดินทางให้นักเรียนสามารถบอกได้ว่าเดินทางไปทิศใดบ้าง เป็นระยะทางเท่าใด	-	-	1	-	1
บทที่ 7 เศษส่วน					
1. เมื่อกำหนดเศษส่วนให้ นักเรียนสามารถเขียนเป็นเศษส่วนที่มีค่าเท่าเดิมโดยที่ตัวเศษหรือตัวส่วนมีค่าตามที่กำหนดให้ได้	1	-	-	-	1
2. เมื่อกำหนดเศษส่วนสองจำนวนให้นักเรียนสามารถเปรียบเทียบและใช้สัญลักษณ์ $>$ $<$ หรือ $=$ ได้	-	1	-	-	1
3. เมื่อกำหนดเศษส่วนใด ๆ ให้นักเรียนสามารถ บวก ลบ คูณ หาร ระคนได้	-	-	1	-	1
4. เมื่อกำหนดเศษส่วนอย่างง่ายที่เฉพาะตัวเศษหรือตัวส่วนที่มีการบวก ลบ คูณ หารอย่างใดอย่างหนึ่งให้ นักเรียนสามารถหาคำตอบได้	-	-	2	-	2
5. เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับการบวก ลบ คูณ หาร เศษส่วนให้ นักเรียนสามารถหาคำตอบได้	-	2	3	1	6

ตาราง 7 (ต่อ)

จุดประสงค์เชิงพฤติกรรม	พฤติกรรมการเรียนรู้				รวม จำนวน ข้อ
	การคิด คำนวณ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
บทที่ 8 การบวก ลบ และคูณทศนิยม					
1. เมื่อกำหนดทศนิยมให้นักเรียนสามารถอ่านและเขียนในรูปกระจายตามค่าประจำหลักได้	1	-	-	-	1
2. เมื่อกำหนดทศนิยมไม่เกิน 3 ตำแหน่งตั้งแต่ 2 จำนวนขึ้นไปให้นักเรียนสามารถเปรียบเทียบและเรียงลำดับจากทศนิยมที่มีค่าน้อยไปหาทศนิยมที่มีค่ามากหรือจากทศนิยมที่มีค่ามากไปทศนิยมที่มีค่าน้อยได้	-	1	-	-	1
3. เมื่อกำหนดทศนิยมให้นักเรียนสามารถเขียนในรูปเศษส่วนได้ หรือเมื่อกำหนดเศษส่วนที่มีตัวส่วนเป็น 10, 100 และ 1,000 ให้นักเรียนสามารถเขียนในรูปทศนิยมได้	1	-	-	-	1
4. เมื่อกำหนดโจทย์เกี่ยวกับการบวก การลบ ทศนิยมที่มีผลลัพธ์ไม่เกิน 3 ตำแหน่งให้นักเรียนสามารถหาคำตอบได้	-	-	1	-	1
5. เมื่อกำหนดจำนวนนับกับทศนิยมหรือทศนิยมกับทศนิยมที่ผลคูณเป็นทศนิยมไม่เกิน 3 ตำแหน่งให้นักเรียนสามารถหาผลคูณได้	-	1	-	-	1
6. เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับการคูณทศนิยมที่ผลคูณเป็นทศนิยมไม่เกิน 3 ตำแหน่งให้นักเรียนสามารถหาคำตอบได้	-	1	3	1	5
รวมจำนวนข้อ	11	18	25	6	60

4. สร้างแบบทดสอบปรนัยแบบเลือกตอบชนิด 4 ตัวเลือก จำนวน 164 ข้อโดยสร้างเกินไว้ประมาณ 2.7 เท่า เพื่อคัดเลือกข้อที่เป็นไปตามเกณฑ์ไว้ใช้ในการวิจัยจำนวน 60 ข้อ

5. ให้ผู้เชี่ยวชาญจำนวน 5 ท่านตรวจสอบแบบทดสอบเป็นรายข้อว่า วัดได้ตรงตามพฤติกรรมการเรียนรู้หรือไม่ โดยใช้เกณฑ์การพิจารณาความเห็นของผู้เชี่ยวชาญ (ค่า IOC) ตั้งแต่ .60 ขึ้นไปถือว่าเป็นข้อสอบที่วัดได้ตรงตามพฤติกรรม (Hambleton.1978 : 34) นำไปใช้ทดสอบได้ ซึ่งผลจากการตรวจสอบปรากฏว่าแบบทดสอบมีค่า IOC ตั้งแต่ .60 - 1 อยู่ในเกณฑ์ที่ใช้ได้ทั้งหมด แต่ต้องนำมาปรับปรุงแก้ไขในเรื่องรูปแบบการพิมพ์ และการใช้ถ้อยคำให้รัดกุมเหมาะสมยิ่งขึ้นในบางข้อ จากนั้นผู้วิจัยได้คัดเลือกข้อสอบไว้จำนวน 60 ข้อเพื่อนำไปใช้ในการทดสอบค่าความยากกับนักเรียนกลุ่มที่สุ่มมาสำหรับการสร้างเครื่องมือ

ตัวอย่าง แบบประเมินความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

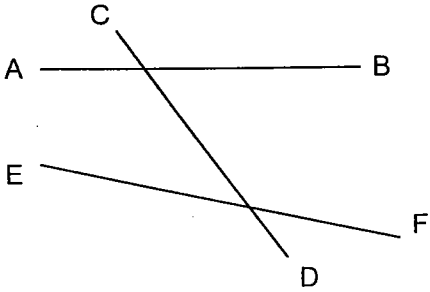
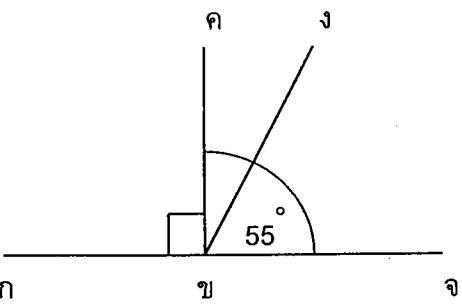
คำชี้แจง ให้ท่านพิจารณาว่า ข้อคำถามแต่ละข้อต่อไปนี้วัดได้ตรงตามพฤติกรรมการเรียนรู้ของวิไลสันจริงหรือไม่โดยให้พิจารณาน้ำหนัก แล้วใส่เครื่องหมาย ✓ ลงในช่องคะแนน

- + 1 เมื่อท่านพิจารณาแล้วเห็นว่า ข้อคำถามข้อนี้วัดได้ตรงตามพฤติกรรมข้อนั้นจริง
- 0 เมื่อท่านพิจารณาแล้วไม่แน่ใจว่า ข้อคำถามข้อนี้วัดได้ตรงตามพฤติกรรมข้อนั้นจริง
- 1 เมื่อท่านพิจารณาแล้วเห็นว่า ข้อคำถามข้อนี้วัดได้ไม่ตรงตามพฤติกรรมข้อนั้นจริง

ตัวอย่าง การประเมินความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

พฤติกรรมการเรียนรู้ที่ต้องการวัด	ข้อสอบ	คะแนน		
		+1	0	-1
การคิดคำนวณ	1. $19026 + (2184 + 593) = \square$ ¹ ก. 21793 ข. 21803 ค. 22503 ง. 22703	✓		

ตัวอย่าง (ต่อ)

พฤติกรรมการเรียนรู้ที่ ต้องการวัด	ข้อสอบ	คะแนน		
		+1	0	-1
ความเข้าใจ	 2. จากรูป มีมุมป้านกี่มุม ก. 4 มุม ข. 6 มุม ค. 8 มุม ง. 10 มุม	✓		
การนำไปใช้	 3. มุม คขง มีขนาดเท่าใด ก. 25° ข. 35° ค. 90° ง. 180°	✓		

ตัวอย่าง (ต่อ)

พฤติกรรมการเรียนรู้ที่ ต้องการวัด	ข้อสอบ	คะแนน		
		+1	0	-1
การวิเคราะห์	ถ้ากำหนดเงินรางวัลให้แก่ผู้แข่งขัน รางวัลที่ 1 เป็นเงิน 25,000 บาท รางวัลที่ 2 เป็นเงิน 10,080 บาท รางวัลที่ 3 เป็นเงิน 7,750 บาท แต่ปรากฏว่ามีผู้ได้รับรางวัลที่ 2 ถึง 2 ราย ส่วนรางวัลที่ 3 ไม่มีผู้ได้รับจะต้องหาเงิน รางวัลเพิ่มอีกกี่บาท ก. 2,330 บาท ข. 19,780 บาท ค. 27,695 บาท ง. 30,025 บาท	✓		

6. นำแบบทดสอบที่คัดเลือกไว้ จำนวน 60 ข้อ ไปทดสอบกับนักเรียนกลุ่มที่สุ่มมา สำหรับการสร้างเครื่องมือ จำนวน 375 คน ตรวจให้คะแนนแบบ 0 - 1 แล้วนำไปวิเคราะห์แบบ ราชคิโมเดล โดยใช้โปรแกรม BICAL เพื่อกำหนดค่าความยาก (b) ผลการวิเคราะห์พบว่า ข้อสอบมีค่าความยากตั้งแต่ -4.57 ถึง +1.93 จึงคัดเลือกข้อสอบที่มีค่าความยากตั้งแต่ -1.9 ถึง +1.9 ไว้จำนวน 50 ข้อ

7. จัดเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

8. จัดพิมพ์เป็นแบบทดสอบเพื่อใช้ในการวิจัย

วิธีการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลที่ใช้ในการศึกษาวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการระหว่างเดือน ธันวาคม 2543 ถึงเดือนมกราคม 2544 โดยทำตามขั้นตอน ดังนี้

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความ อนุเคราะห์ในการเก็บข้อมูล

2. ติดต่อกับโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนดวัน เวลา ในการสอบ

3. เตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่สอบในแต่ละครั้ง
4. นำแบบทดสอบวัดผลสัมฤทธิ์ที่สร้างขึ้นไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง
5. อธิบายให้นักเรียนที่เป็นกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ และประโยชน์ที่ได้รับจากการทำแบบทดสอบ
6. ตรวจสอบกระดาษคำตอบตามวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ

วิธีการตรวจให้คะแนนโดยกฎชิลลิ่ง 3 แบบ

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง
2. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง
3. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง

ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 แล้วกำหนดการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด ซึ่งมีการให้คะแนนในแต่ละตำแหน่ง ดังนี้

1.1 การตอบผิดตำแหน่งที่หนึ่งให้ 1 คะแนนสำหรับข้อที่ตอบถูกตั้งแต่ข้อที่ 1 ถึง ข้อที่ตอบผิด (0) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.2 การตอบผิดตำแหน่งที่สองให้ 1 คะแนนสำหรับข้อที่ตอบถูกตั้งแต่ข้อที่ 1 ถึง ข้อที่ตอบผิดติดต่อกันตั้งแต่สองข้อ (00) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.3 การตอบผิดตำแหน่งที่สามให้ 1 คะแนนสำหรับข้อที่ตอบถูกตั้งแต่ข้อที่ 1 ถึง ข้อที่ตอบผิดติดต่อกันตั้งแต่สามข้อ (000) เป็นครั้งแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

1.4 การไม่กำหนดตำแหน่งการตอบผิด ตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

ตัวอย่าง การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง

คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1
1.	1	14.	0	27.	1	40.	1
2.	0*	15.	1	28.	1	41.	1
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.	0***
9.	0**	22.	0	35.	1	48.	0
10.	1	23.	1	36.	0	49.	1
11.	1	24.	1	37.	1	50.	1
12.	1	25.	0	38.	1		
13.	1	26.	1	39.	0		

หมายเหตุ

* การตอบผิดตำแหน่งที่หนึ่ง นักเรียนจะได้คะแนน 1 คะแนน

** การตอบผิดตำแหน่งที่สอง นักเรียนจะได้คะแนน 4 คะแนน

*** การตอบผิดตำแหน่งที่สาม นักเรียนจะได้คะแนน 28 คะแนน

การไม่กำหนดตำแหน่งการตอบผิด นักเรียนจะได้คะแนน 30 คะแนน

1

ผลจากการกำหนดคะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องจะได้คะแนน 4 ชุด คือ

ชุดที่ 1 คะแนนจากการตอบผิดตำแหน่งที่หนึ่ง

ชุดที่ 2 คะแนนจากการตอบผิดตำแหน่งที่สอง

ชุดที่ 3 คะแนนจากการตอบผิดตำแหน่งที่สาม

ชุดที่ 4 คะแนนจากการไม่กำหนดตำแหน่งการตอบผิด

2. การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อหนึ่งเพียงตัวเดียว หลาย ๆ ครั้ง

ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 แล้วกำหนดการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งการตอบผิด ซึ่งมีการตรวจให้คะแนนในแต่ละตำแหน่ง ดังนี้

2.1 การตอบผิดตำแหน่งที่หนึ่ง ให้ 1 คะแนนตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิด (0) เป็นข้อแรก ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.2 การตอบผิดตำแหน่งที่สอง ให้ 1 คะแนนตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดเป็นข้อที่สอง (0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.3 การตอบผิดตำแหน่งที่สาม ให้ 1 คะแนนตั้งแต่ข้อที่ 1 ถึงข้อที่ตอบผิดเป็นข้อที่สาม (0,0,0) ข้อที่ตอบถูกหลังจากนั้นจะไม่นำมานับคะแนน

2.4 การไม่กำหนดตำแหน่งการตอบผิด ตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

ตัวอย่าง การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง

คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1
1.	1	14.	0	27.	1	40.	1
2.	0*	15.	1	28.	1	41.	1
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.	0
9.	0	22.	0	35.	1	48.	0
10.	1	23.	1	36.	0	49.	1
11.	1	24.	1	37.	1	50.	1
12.	1	25.	0	38.	1		
13.	1	26.	1	39.	0		

หมายเหตุ

- * การตอบผิดตำแหน่งที่หนึ่ง นักเรียนจะได้คะแนน 1 คะแนน
- ** การตอบผิดตำแหน่งที่สอง นักเรียนจะได้คะแนน 2 คะแนน
- *** การตอบผิดตำแหน่งที่สาม นักเรียนจะได้คะแนน 3 คะแนน
- การไม่กำหนดตำแหน่งการตอบผิด นักเรียนจะได้คะแนน 30 คะแนน

ผลจากการตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียว
หลาย ๆ ครั้ง จะได้คะแนน 4 ชุด คือ

- ชุดที่ 1 คะแนนจากการตอบผิดตำแหน่งที่หนึ่ง
- ชุดที่ 2 คะแนนจากการตอบผิดตำแหน่งที่สอง
- ชุดที่ 3 คะแนนจากการตอบผิดตำแหน่งที่สาม
- ชุดที่ 4 คะแนนจากการไม่กำหนดตำแหน่งการตอบผิด

3. การตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อหนึ่งมากกว่า 1 ตัว

ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 แล้วกำหนดการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม และการไม่กำหนดตำแหน่งที่การตอบผิด ซึ่งมีการตรวจให้คะแนนในแต่ละตำแหน่ง ดังนี้

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

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

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

3.4 การไม่กำหนดตำแหน่งการตอบผิด ตรวจให้คะแนนแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

ตัวอย่าง การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว

คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1	คนที่ ข้อที่	1
1.	1	14.	0	27.	1	40.	1
2.	0	15.	1	28.	1	41.	1
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.	0
9.	0*	22.	0	35.	1	48.	0
10.	1	23.	1	36.	0	49.	1
11.	1	24.	1	37.	1	50.	1
12.	1	25.	0	38.	1		
13.	1	26.	1	39.	0		

หมายเหตุ

- * การตอบผิดตำแหน่งที่หนึ่ง นักเรียนจะได้คะแนน 4 คะแนน
- ** การตอบผิดตำแหน่งที่สอง นักเรียนจะได้คะแนน 19 คะแนน
- *** การตอบผิดตำแหน่งที่สาม นักเรียนจะได้คะแนน 28 คะแนน
- การไม่กำหนดตำแหน่งที่การตอบผิด นักเรียนจะได้คะแนน 30 คะแนน

ผลจากการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จะได้คะแนน 4 ชุด คือ

- ชุดที่ 1 คะแนนจากการตอบผิดตำแหน่งที่หนึ่ง
- ชุดที่ 2 คะแนนจากการตอบผิดตำแหน่งที่สอง
- ชุดที่ 3 คะแนนจากการตอบผิดตำแหน่งที่สาม
- ชุดที่ 4 คะแนนจากการไม่กำหนดตำแหน่งการตอบผิด

ขั้นตอนในการวิเคราะห์ข้อมูล

ขั้นตอนในการวิเคราะห์ข้อมูลมีการดำเนินการตามลำดับดังต่อไปนี้

1. การวิเคราะห์ข้อมูลในการทดลองเครื่องมือ (Try Out) ผู้วิจัยได้ดำเนินการดังนี้

1.1 ตรวจให้คะแนนแบบทดสอบแบบ 0 - 1 ตั้งแต่ข้อที่ 1 จนถึงข้อสุดท้าย

1.2 วิเคราะห์ค่าความยากด้วยวิธีโมเดล (Rasch Model) โปรแกรม BICAL

เพื่อจัดเรียงข้อสอบตามค่าความยากโดยเรียงจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด

2. การวิเคราะห์ข้อมูลในการวิจัย ผู้วิจัยได้ดำเนินการดังนี้

2.1 ตรวจให้คะแนนแบบทดสอบจำนวน 790 ฉบับ โดยกฎข้อที่แตกต่างกันทั้ง 3 แบบและนำเฉพาะข้อมูลที่ตรวจได้ครบถึงการตอบผิดตำแหน่งที่ 3 จำนวน 755 ฉบับ มาสุ่มแบบใส่คืนโดยเครื่องคอมพิวเตอร์โปรแกรม SPSS เป็นกลุ่ม ๆ ละ 50 คน จำนวน 14 กลุ่ม

2.2 หาค่าสถิติพื้นฐาน คือ

2.2.1 คะแนนเฉลี่ย (Mean)

2.2.2 ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

2.2.3 สัมประสิทธิ์การกระจาย (Garrett. 1966 : 57)

2.3 ประเมินค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตรสัมประสิทธิ์ r_B (บุญเชิด ภิญโญนันตพงษ์. 2538 : 51) แล้วแปลงค่าความเชื่อมั่นเป็นคะแนนมาตรฐานซี (Fisher's Z Transformation)

3. การทดสอบสมมติฐาน

3.1 เปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎข้อที่แบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน ในสมมติฐานข้อที่ 1 โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

3.2 เปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎข้อที่แบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน ในสมมติฐานข้อที่ 2 โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

3.3 เปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎข้อที่แบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน ในสมมติฐานข้อที่ 3 โดย

วิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

3.4 เปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกันในสมมติฐานข้อที่ 4 โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

3.5 เปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สองมีค่าแตกต่างกันในสมมติฐานข้อที่ 5 โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

3.6 เปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกัน ในสมมติฐานข้อที่ 6 โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

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

2. วิเคราะห์ข้อสอบรายข้อเพื่อหาค่าความยากด้วยราชคโมเดล (Rasch Model) โปรแกรม BICAL

3. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่ง โดยใช้สูตรสัมประสิทธิ์ r_b (บุญเชิด ภิญโญนนตพงษ์. 2538 : 51)

4. สถิติทดสอบนัยสำคัญความแตกต่างค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซึลลิ่งแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

5. สถิติทดสอบนัยสำคัญความแตกต่างของค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งทั้ง 3 แบบจากการตอบผิดตำแหน่งเดียวกันของแบบทดสอบ โดยวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) และทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's test) (Scheffe'. 1959)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้เข้าใจตรงกันในการแปลความหมายผลของการวิเคราะห์ข้อมูล ผู้วิจัยจึงกำหนดสัญลักษณ์ และอักษรย่อสำหรับใช้ในการวิเคราะห์ข้อมูล ดังนี้

r_B	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
$\bar{r}_B$	แทน	ค่าเฉลี่ยความเชื่อมั่นของแบบทดสอบ
df	แทน	ขั้นของความเป็นอิสระ (Degrees of Freedom)
F	แทน	ค่าสถิติที่ใช้ในการพิจารณา (F - Distribution)
$\bar{Z}(r_B)$	แทน	คะแนนมาตรฐาน Z ที่แปลงจากค่า r_B (Fisher's Z Transformation)
แบบที่ 1	แทน	กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่อง
แบบที่ 2	แทน	กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง
แบบที่ 3	แทน	กฎซัลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว
ด 1	แทน	การตอบผิดตำแหน่งที่หนึ่ง
ด 2	แทน	การตอบผิดตำแหน่งที่สอง
ด 3	แทน	การตอบผิดตำแหน่งที่สาม
*	แทน	มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลนำเสนอเป็น 4 ตอน คือ

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

ตอนที่ 2 การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎซัลลิ่งที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่ง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซัลลิ่งแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

3.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

3.2 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

3.3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งเดียวกันของแบบทดสอบ

4.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบจากการตอบผิดตำแหน่งที่หนึ่ง เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

4.2 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สอง เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

4.3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สาม เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบ

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนผลการสอบของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่เรียงตามค่าความยากจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด จำนวน 50 ข้อที่ตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด ของกลุ่มตัวอย่างจำนวน 14 กลุ่มมาวิเคราะห์หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์การกระจาย (CV) ปรากฏผลดังแสดงในตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

วิธีการตรวจให้คะแนน ตำแหน่งการตอบผิด	กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง			กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง			กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว		
	$\bar{X}$	S.D.	CV	$\bar{X}$	S.D.	CV	$\bar{X}$	S.D.	CV
ตำแหน่งที่หนึ่ง	7.76	6.32	81.44	7.76	6.32	81.44	16.20	9.37	57.83
ตำแหน่งที่สอง	16.20	9.37	57.83	11.88	7.46	62.79	19.76	9.29	47.01
ตำแหน่งที่สาม	20.48	9.21	44.97	14.39	7.80	54.20	21.26	9.19	43.22
การไม่กำหนดตำแหน่ง	27.97	6.48	23.16	27.97	6.48	23.16	27.97	6.48	23.16

ผลการวิเคราะห์ตามตาราง 8 ปรากฏว่า

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 7.76 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุด และการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 81.44 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

2. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยตั้งแต่ 7.76 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่ง

การตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 81.44 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุดและการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

3. การตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าเฉลี่ยตั้งแต่ 16.20 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 57.83 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุด และการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

ตอนที่ 2 การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิกสิงที่แตกต่างกันทั้ง 3 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนผลการสอบของแบบทดสอบมาตรวจให้คะแนนโดยกฎซิกสิงทั้ง 3 แบบจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมาหาค่าความยากแต่ละข้อ ค่าความยากทั้งฉบับ สัดส่วนของผู้ตอบข้อสอบสองข้อใด ๆ ถูกทั้งสองข้อ ค่าแลมด้า (λ) ค่าความแปรปรวนของคะแนนทั้งฉบับ แล้วนำค่าที่ได้ ไปแทนค่าในสูตรการหาค่าสัมประสิทธิ์ r_B (บุญเชิด ภิบุญญ อนันตพงษ์. 2538 : 51) เพื่อหาค่าความเชื่อมั่นของแบบทดสอบ จากนั้นจึงแปลงค่าความเชื่อมั่นเป็นคะแนนมาตรฐานซี (Fisher's Z Transformation) ปรากฏผลดังแสดงในตาราง 9

ตาราง 9 ค่าความเชื่อมั่น และค่าที่แปลงเป็นคะแนนมาตรฐานซี ของแบบทดสอบที่ตรวจให้
คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่
กำหนดตำแหน่งการตอบผิด

ค่าความ เชื่อมั่น กลุ่มที่		กฎซึลลิ่งแบบคะแนน ศูนย์ต่อเนื่อง			กฎซึลลิ่งแบบคะแนน ศูนย์ต่อเนื่องเพียง ตัวเดียวหลาย ๆ ครั้ง			กฎซึลลิ่งแบบคะแนน ศูนย์ต่อเนื่องมากกว่า 1 ตัว			การไม่ กำหนด ตำแหน่ง การตอบ ผิด
		ต 1	ต 2	ต 3	ต 1	ต 2	ต 3	ต 1	ต 2	ต 3	
1.	r_B	.96	.95	.92	.96	.95	.94	.95	.93	.93	.71
	$Z(r_B)$	1.94	1.88	1.65	1.94	1.83	1.78	1.88	1.69	1.65	.89
2.	r_B	.96	.96	.94	.96	.95	.95	.96	.94	.94	.79
	$Z(r_B)$	2.01	1.94	1.73	2.01	1.94	1.88	1.94	1.78	1.73	1.09
3.	r_B	.95	.96	.93	.95	.95	.95	.96	.94	.92	.76
	$Z(r_B)$	1.88	1.94	1.69	1.88	1.88	1.83	1.94	1.78	1.65	1.00
4.	r_B	.95	.95	.94	.95	.94	.94	.95	.94	.94	.73
	$Z(r_B)$	1.94	1.83	1.78	1.94	1.78	1.78	1.83	1.78	1.73	.95
5.	r_B	.96	.95	.94	.96	.95	.94	.95	.92	.91	.72
	$Z(r_B)$	1.94	1.88	1.73	1.94	1.83	1.78	1.88	1.58	1.55	.91
6.	r_B	.96	.96	.96	.96	.96	.95	.96	.96	.95	.82
	$Z(r_B)$	1.94	2.01	1.94	1.94	1.94	1.94	2.01	1.94	1.88	1.15
7.	r_B	.95	.94	.91	.95	.94	.94	.94	.93	.92	.76
	$Z(r_B)$	1.88	1.83	1.55	1.88	1.83	1.73	1.83	1.69	1.62	1.00
8.	r_B	.96	.96	.94	.96	.96	.95	.96	.95	.94	.81
	$Z(r_B)$	2.01	2.01	1.78	2.01	2.01	1.88	2.01	1.83	1.83	1.14
9.	r_B	.96	.96	.94	.96	.95	.95	.96	.95	.94	.85
	$Z(r_B)$	2.01	1.94	1.78	2.01	1.88	1.88	1.94	1.83	1.73	1.25
10.	r_B	.97	.95	.94	.97	.95	.95	.95	.94	.93	.79
	$Z(r_B)$	2.09	1.88	1.78	2.09	1.94	1.88	1.88	1.78	1.65	1.07
11.	r_B	.96	.96	.94	.96	.96	.95	.96	.94	.94	.79
	$Z(r_B)$	1.94	1.94	1.78	1.94	1.94	1.88	1.94	1.78	1.73	1.08

ตาราง 9 (ต่อ)

ค่าความ เชื่อมั่น	กลุ่มที่	กฎชิลลิ่งแบบคะแนน ศูนย์ต่อเนื่อง			กฎชิลลิ่งแบบคะแนน ศูนย์ต่อเนื่องเพียง ตัวเดียวหลาย ๆ ครั้ง			กฎชิลลิ่งแบบคะแนน ศูนย์ต่อเนื่องมากกว่า 1 ตัว			การไม่ กำหนด ตำแหน่ง การตอบ ผิด
		ต 1	ต 2	ต 3	ต 1	ต 2	ต 3	ต 1	ต 2	ต 3	
	r_B	.96	.96	.93	.96	.96	.95	.96	.94	.93	.80
12.	$Z(r_B)$	1.94	1.94	1.69	1.94	2.01	1.88	1.94	1.78	1.73	1.11
	r_B	.96	.95	.93	.96	.96	.94	.95	.94	.93	.78
13.	$Z(r_B)$	2.01	1.94	1.69	2.01	1.94	1.83	1.94	1.78	1.65	1.07
	r_B	.95	.95	.92	.95	.95	.94	.95	.93	.92	.78
14.	$Z(r_B)$	1.94	1.88	1.62	1.94	1.88	1.78	1.88	1.69	1.62	1.05
	$\bar{r}_B$	.96	.95	.93	.96	.95	.95	.95	.94	.93	.78
	$\bar{Z}(r_B)$	1.96	1.92	1.73	1.96	1.90	1.84	1.92	1.76	1.70	1.05

ผลการวิเคราะห์ตามตาราง 9 ปรากฏว่า

1. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สามมีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .93 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

2. การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .95 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

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

4. การตรวจให้คะแนนโดยไม่กำหนดตำแหน่งการตอบผิดมีค่าความเชื่อมั่นเฉลี่ยต่ำสุดเท่ากับ .78

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

การวิเคราะห์ตอนนี้ ผู้วิจัยนำค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิดที่แปลงเป็นคะแนนมาตรฐานขึ้นมาเปรียบเทียบ และทดสอบความแตกต่าง ดังนี้

3.1 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิดมาเปรียบเทียบ โดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 10

ตาราง 10 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

	ตำแหน่งที่ หนึ่ง	ตำแหน่งที่ สอง	ตำแหน่งที่ สาม	การไม่ กำหนด ตำแหน่ง	F	df
$\bar{Z}(r_B)$	1.967	1.922	1.733	1.059	822.330 *	3,39
$S(r_B)$	.055	.056	.092	.097		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 10 ปรากฏว่าค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดอย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้น จึงทดสอบความแตกต่าง

ค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe', 1959) ปรากฏผลดังแสดงในตาราง 11

ตาราง 11 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฏชิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

$\bar{Z} (r_B)$	ตำแหน่งการตอบผิด	การไม่กำหนดตำแหน่ง	ตำแหน่งที่สาม	ตำแหน่งที่สอง	ตำแหน่งที่หนึ่ง
1.0595	การไม่กำหนดตำแหน่ง	-	.6738 *	.8628 *	.9078 *
1.7333	ตำแหน่งที่สาม		-	.1890 *	.2340 *
1.9223	ตำแหน่งที่สอง			-	.0450
1.9673	ตำแหน่งที่หนึ่ง				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 11 ปรากฏว่าค่าความเชื่อมั่นรายคู่ของการตรวจให้คะแนนโดยกฏชิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เกือบทั้งหมด ยกเว้นการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สองที่มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.2 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฏชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมาเปรียบเทียบ โดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 12

ตาราง 12 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนน โดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

	ตำแหน่งที่ หนึ่ง	ตำแหน่งที่ สอง	ตำแหน่งที่ สาม	การไม่ กำหนด ตำแหน่ง	F	df
$\bar{Z}(r_B)$	1.967	1.906	1.842	1.059	1271 *	3,39
$S(r_B)$	.055	.070	.060	.097		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 12 ปรากฏว่าค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด ที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดอย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้นจึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe'. 1959) ปรากฏผลดังแสดงในตาราง 13

ตาราง 13 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

$\bar{Z} (r_b)$	ตำแหน่งการตอบผิด	การไม่กำหนดตำแหน่ง	ตำแหน่งที่สาม	ตำแหน่งที่สอง	ตำแหน่งที่หนึ่ง
1.0595	การไม่กำหนดตำแหน่ง	-	.7831 *	.8473 *	.9078 *
1.8426	ตำแหน่งที่สาม		-	.0642	.1247 *
1.9068	ตำแหน่งที่สอง			-	.0605
1.9673	ตำแหน่งที่หนึ่ง				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 13 ปรากฏว่าค่าความเชื่อมั่นรายคู่ของการตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เกือบทั้งหมด ยกเว้นการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง และการตอบผิดตำแหน่งที่สองกับตำแหน่งที่สาม ที่มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.3 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) ผู้วิจัยนำค่าความเชื่อมั่นการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมาเปรียบเทียบ โดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 14

ตาราง 14 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนน โดยกฎซัลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

	ตำแหน่งที่ หนึ่ง	ตำแหน่งที่ สอง	ตำแหน่งที่ สาม	การไม่ กำหนด ตำแหน่ง	F	df
$\bar{Z}(r_B)$	1.922	1.769	1.703	1.059	1014 *	3,39
$S(r_B)$	.056	.082	.087	.097		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 14 ปรากฏว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซัลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิด ที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ค่าความเชื่อมั่นของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดอย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้น จึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe', 1959) ปรากฏผลดังแสดงในตาราง 15

ตาราง 15 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎ
ซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดต่างกัน 3
ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด

$\bar{Z} (r_B)$	ตำแหน่งการตอบผิด	การไม่ กำหนด ตำแหน่ง	ตำแหน่ง ที่ สาม	ตำแหน่ง ที่ สอง	ตำแหน่ง ที่ หนึ่ง
1.0595	การไม่กำหนดตำแหน่ง	-	.6436 *	.7099 *	.8628 *
1.7031	ตำแหน่งที่สาม		-	.0663	.2192 *
1.7694	ตำแหน่งที่สอง			-	.1529 *
1.9223	ตำแหน่งที่หนึ่ง				-

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 15 ปรากฏว่าค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เกือบทั้งหมด ยกเว้นการตอบผิดตำแหน่งที่สองกับตำแหน่งที่สามที่มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการกำหนดการตอบผิดตำแหน่งเดียวกัน

การวิเคราะห์ตอนนี้ ผู้วิจัยนำค่าของความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบ จากการตอบผิดตำแหน่งเดียวกันที่แปลงเป็นคะแนนมาตรฐานขึ้นมาเปรียบเทียบ และทดสอบความแตกต่าง ดังนี้

4.1 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบ ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งของการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบมาเปรียบเทียบ โดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 16

ตาราง 16 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบ จากการตอบผิดตำแหน่งที่หนึ่ง

	แบบที่ 1	แบบที่ 2	แบบที่ 3	F	df
$\bar{Z}(r_B)$	1.967	1.967	1.922	4.500 *	2,26
$S(r_B)$	.055	.055	.056		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 16 ปรากฏว่าค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งของการตรวจให้คะแนนโดยกฎซึลลิ่งทั้ง 3 แบบที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่อง กฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และกฎซึลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว อย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้นจึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe'. 1959) ปรากฏผลดังแสดงในตาราง 17

ตาราง 17 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎซึลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง

$\bar{Z}(r_B)$	การตรวจให้คะแนนโดยกฎซึลลิ่ง	แบบที่ 3	แบบที่ 1	แบบที่ 2
1.9223	แบบที่ 3	-	.0450	.0450
1.9673	แบบที่ 1		-	.0000
1.9673	แบบที่ 2			-

ผลการวิเคราะห์ตามตาราง 17 ปรากฏว่า ค่าความเชื่อมั่นรายคู่ จากการตอบผิดตำแหน่งที่หนึ่งของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.2 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบ ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สองของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ มาเปรียบเทียบโดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 18

ตาราง 18 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สอง

	แบบที่ 1	แบบที่ 2	แบบที่ 3	F	df
$\bar{Z}(r_B)$	1.922	1.906	1.769	49.500 *	2,26
$S(r_B)$	.056	.070	.082		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 18 ปรากฏว่า ค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สองของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง กฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการตอบผิดตำแหน่งที่สองอย่างน้อย 1 คู่ มีค่าแตกต่างกัน ดังนั้น จึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe'. 1959) ปรากฏผลดังแสดงในตาราง 19

ตาราง 19 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สอง

$\bar{Z} (r_B)$	การตรวจให้คะแนนโดยกฎชิลลิ่ง	แบบที่ 3	แบบที่ 2	แบบที่ 1
1.7694	แบบที่ 3	-	.1374 *	.1529 *
1.9068	แบบที่ 2		-	.0155
1.9223	แบบที่ 1			-

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 19 ปรากฏว่า ค่าความเชื่อมั่นรายคู่จากการตอบผิดตำแหน่งที่สองของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เกือบทั้งหมด ยกเว้นการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) กับการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) ที่มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.3 ในการวิเคราะห์เพื่อเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบ ผู้วิจัยนำค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สามของการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบ มาเปรียบเทียบ โดยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ Repeated Measures ANOVA (Winer. 1971 : 261) ปรากฏผลดังแสดงในตาราง 20

ตาราง 20 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สาม

	แบบที่ 1	แบบที่ 2	แบบที่ 3	F	df
$\bar{Z} (r_B)$	1.733	1.842	1.703	37.500 *	2,26
$S(r_B)$	.092	.060	.087		

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 20 ปรากฏว่า ค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สามของการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบ ที่คำนวณด้วยสูตร r_B มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ค่าความเชื่อมั่นของการตรวจให้คะแนน โดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง กฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว อย่างน้อย 1 คู่มีค่าแตกต่างกัน ดังนั้นจึงทดสอบความแตกต่างค่าความเชื่อมั่นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) (Scheffe', 1959) ปรากฏผลดังแสดงในตาราง 21

ตาราง 21 การเปรียบเทียบค่าความเชื่อมั่นรายคู่ระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สาม

$\bar{Z} (r_B)$	การตรวจให้คะแนนโดยกฎซิลลิง	แบบที่ 3	แบบที่ 1	แบบที่ 2
1.7031	แบบที่ 3	-	.0302	.1395 *
1.7333	แบบที่ 1		-	.1093 *
1.8426	แบบที่ 2			-

* มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 21 ปรากฏว่า ค่าความเชื่อมั่นรายคู่จากการตอบผิดตำแหน่งที่สามของการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบ มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เกือบทั้งหมดยกเว้นการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) ที่มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

การศึกษาวิจัยครั้งนี้ เป็นการศึกษาค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกัน 3 แบบ จากการตอบผิดต่างกัน 3 ตำแหน่งกับการ ไม่กำหนดตำแหน่งการตอบผิด และค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิง ทั้ง 3 แบบจากการตอบผิดตำแหน่งเดียวกันว่าจะมีค่าความเชื่อมั่นแตกต่างกันหรือไม่ ซึ่งสรุป สาระสำคัญได้ดังนี้

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง
2. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียว หลาย ๆ ครั้ง
3. เพื่อศึกษาความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว
4. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องจาก การกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด
5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง เพียงตัวเดียวหลาย ๆ ครั้ง จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่ กำหนดตำแหน่งการตอบผิด
6. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง มากกว่า 1 ตัว จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนด ตำแหน่งการตอบผิด
7. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎชิลลิงที่แตกต่างกันทั้ง 3 แบบจาก การตอบผิดตำแหน่งที่หนึ่ง

8. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สอง

9. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งที่สาม

สมมติฐานในการศึกษาค้นคว้า

การศึกษาครั้งนี้ได้กำหนดสมมติฐาน ดังนี้

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
2. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งจากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
3. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว จากการกำหนดตำแหน่งการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าแตกต่างกัน
4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกัน
5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สองมีค่าแตกต่างกัน
6. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกัน

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการศึกษา ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ จำนวน 444 โรงเรียน 490 ห้องเรียน และมีนักเรียนจำนวน 9,684 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการศึกษาได้แก่นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ เลือกมาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) สุ่มได้ 35 ห้องเรียน นักเรียนจำนวน 790 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 เนื้อหาบทที่ 1 - 8 แบบปรนัยแบบเลือกตอบชนิด 4 ตัวเลือก 1 ฉบับ จำนวน 50 ข้อที่ผู้วิจัยสร้างขึ้นตามพฤติกรรมการเรียนรู้ของวิลสัน และมีการเรียงข้อคำถามตามค่าความยากโดยเรียงจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด

สรุปผลการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้มุ่งที่จะศึกษาค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแต่ละแบบจากการตอบผิดต่างกัน 3 ตำแหน่งกับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกัน และค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบจากการตอบผิดตำแหน่งเดียวกันมีค่าความเชื่อมั่นแตกต่างกัน ผลการวิจัยมีดังนี้

1. ค่าสถิติพื้นฐาน

1.1 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยตั้งแต่ 7.76 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 81.44 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุดและการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

1.2 การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยตั้งแต่ 7.76 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุด และการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 81.44 โดยสัมประสิทธิ์การกระจายของการตรวจให้

คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุดและการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

1.3 การตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง ตำแหน่งที่สาม กับการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยตั้งแต่ 16.20 - 27.97 โดยการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีค่าเฉลี่ยสูงสุดและการตอบผิดตำแหน่งที่หนึ่งมีค่าเฉลี่ยต่ำสุด ส่วนสัมประสิทธิ์การกระจายมีค่าตั้งแต่ 23.16 - 57.83 โดยสัมประสิทธิ์การกระจายของการตรวจให้คะแนนจากการไม่กำหนดตำแหน่งการตอบผิดมีการกระจายต่ำสุดและการตอบผิดตำแหน่งที่หนึ่งมีการกระจายสูงสุด

2. ค่าความเชื่อมั่น

2.1 การตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .93 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

2.2 การตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สองและตำแหน่งที่สามมีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .95 - .96 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

2.3 การตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ตำแหน่งที่สอง และตำแหน่งที่สาม มีค่าความเชื่อมั่นเฉลี่ยตั้งแต่ .93 - .95 โดยการตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นเฉลี่ยสูงสุด

2.4 การตรวจให้คะแนนโดยไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นเฉลี่ยต่ำสุดเท่ากับ .78

3. การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิกสิงแต่ละแบบ จากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด

3.1 การเปรียบเทียบค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิกสิงแบบคะแนนศูนย์ต่อเนื่องจากการตอบผิดต่างกัน 3 ตำแหน่ง กับการไม่กำหนดตำแหน่งการตอบผิด มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่การตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม และการตอบผิด

โดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง กับการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.3 การเปรียบเทียบค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งที่แตกต่างกันทั้ง 3 แบบของแบบทดสอบ จากการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ การตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องกับการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งกับการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว ส่วนการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องกับการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

จากผลการวิเคราะห์ข้อมูล พบว่า

1. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง (แบบที่ 1) จากการตอบผิดตำแหน่งที่หนึ่ง กับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม และการตอบผิดตำแหน่งที่สองกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และ สอดคล้องกับผลการศึกษาของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 52) วิลสัน และคณะ (Wilson and others. 1988 : 216) จินตนา ต่อทรัพย์สินชัย (2533 : 47) และมานะ คิวเจริญ (2539 : 60) แต่เมื่อพิจารณาถึงความแตกต่างพบว่า มีส่วนที่ต่างกัน 2 ประการ ประการแรก การตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุดต่างกับผลการศึกษาของสมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 52) วิลสัน และ คณะ (Wilson and others. 1988 : 216) ที่พบว่า การตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด เหตุที่เป็นเช่นนี้อาจเนื่องมาจากเหตุผล 3 ข้อ คือ 1) แบบทดสอบที่ผู้วิจัยใช้สร้างขึ้นตามพฤติกรรมการเรียนรู้ของวิลสัน (Wilson. 1971 : 645 - 696) ประกอบด้วยความรู้ด้านการคิดคำนวณ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ ส่วนแบบทดสอบที่สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987.) วิลสัน และคณะ (Wilson and others. 1988.) ใช้เป็นแบบทดสอบมาตรฐานที่แบ่งเป็นฉบับย่อยฉบับละพฤติกรรม 2) จำนวนคนที่ใช้ในการศึกษาวิจัยต่างกัน คือ ผู้วิจัยศึกษากับกลุ่มตัวอย่าง 14 กลุ่ม ๆ ละ 50 คน ส่วนกลุ่มตัวอย่างที่สมบูรณ์ สุริยวงศ์ (Suriyawongse,

Somboon. 1987.) วิลสันและคณะ (Wilson and others. 1988.) ใช้มีจำนวน 7 กลุ่ม ๆ ละ 100 คน (Edward. 1957 ; Torgerson. 1958.) 3) ใช้สูตรการคำนวณค่าความเชื่อมั่นต่างกัน คือ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987.) วิลสัน และคณะ (Wilson and others. 1988.) หาค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์แอลฟา (Alpha Coefficient) ส่วนการศึกษาวิจัยครั้งนี้ผู้วิจัยหาค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์ r_B (Coefficient r_B) (บุญเชิด ภิญโญ อนันตพงษ์. 2538 : 51) และความแตกต่างประการที่สองการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สองมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติอาจเนื่องมาจากข้อสอบมีจำนวนข้อมากถึง 50 ข้อรวมกับความยากที่เพิ่มมากขึ้นทำให้นักเรียนเบื่อ จึงขาดความรอบคอบในการทำแบบทดสอบ

2. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบ ที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง (แบบที่ 2) จากการตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด และการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการศึกษาของสมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 52) วิลสัน และคณะ (Wilson and others. 1988 : 216) จินตนา ต่อทรัพย์สินชัย (2533 : 47) และ มานะ คิ้วเจริญ (2539 : 60) แต่เมื่อพิจารณาถึงความแตกต่างพบว่า การตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุดต่างกับผลการศึกษาของสมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987 : 52) วิลสัน และคณะ (Wilson and others. 1988 : 216) ที่พบว่าการตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด เหตุที่เป็นเช่นนี้อาจเนื่องมาจากเหตุผล 3 ข้อ ดังได้กล่าวไว้ในข้อที่ 1 และอาจเนื่องมาจากเงื่อนไขในการตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่อง กับเงื่อนไขในการตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งต่างกัน และแม้ว่าค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งจะมีค่าสูงสุดต่างกับผลการศึกษาของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987.) วิลสันและคณะ (Wilson and others. 1988.) แต่ก็สอดคล้องกับผลการศึกษาของ มานะ คิ้วเจริญ (2539 : 60) อาจเป็นเพราะแบบทดสอบที่ใช้สร้างตามพฤติกรรมการเรียนรู้ของวิลสัน (Wilson. 1971 : 645 - 696) และมีเงื่อนไขในการตรวจให้คะแนนแบบเดียวกัน ส่วนการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สอง มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ อาจเป็นเพราะเงื่อนไขในการตรวจให้คะแนนที่เปิดโอกาสให้ตอบผิดได้เพียงครั้งเดียว

3. ค่าความเชื่อมั่นระหว่างตำแหน่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว (แบบที่ 3) จากการตอบผิดตำแหน่งที่หนึ่งกับการไม่

กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สองกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่สามกับการไม่กำหนดตำแหน่งการตอบผิด การตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง และการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สาม มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการศึกษาของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987.) วิลสัน และ คณะ (Wilson and others. 1988.) จินตนา ต่อทรัพย์สินชัย (2533 : 47) และ มานะ คิ้วเจริญ (2539 : 60) แต่เมื่อพิจารณาถึงความแตกต่างพบว่า การตอบผิดตำแหน่งที่หนึ่งมีค่าความเชื่อมั่นสูงสุด ต่างกับผลการศึกษาวิจัยของ สมบูรณ์ สุริยวงศ์ (Suriyawongse, Somboon. 1987.) วิลสัน และคณะ (Wilson and others. 1988.) ที่พบว่า การตอบผิดตำแหน่งที่สองมีค่าความเชื่อมั่นสูงสุด ข้อแตกต่างดังได้กล่าวนี้อาจเกิดจากเงื่อนไขการตรวจให้คะแนน คือการตอบผิดตำแหน่งที่หนึ่งโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว ก็คือตำแหน่งเดียวกันกับการตอบผิดตำแหน่งที่สองของกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องนั่นเอง และจากการพิจารณาเราจะเห็นได้ว่าการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว มีค่าความเชื่อมั่นระหว่างตำแหน่งแตกต่างกันในทุกตำแหน่ง คือการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง การตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สาม การตอบผิดตำแหน่งที่หนึ่งกับการไม่กำหนดตำแหน่งการตอบผิด และแตกต่างจากวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องและกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง ในการศึกษาวิจัยครั้งนี้ผู้วิจัยมีความเห็นว่าวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว น่าจะเป็นวิธีการตรวจให้คะแนนที่ให้ผลดีกว่าวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งสองแบบแรก

4. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบของแบบทดสอบจากการตอบผิดตำแหน่งที่หนึ่ง มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ เมื่อทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเชฟเฟ (Scheffe's Test) ปรากฏว่าไม่พบความแตกต่างอาจเนื่องมาจากเหตุผล 2 ประการ คือ 1) ค่าความเชื่อมั่นของแบบทดสอบที่ตรวจให้คะแนนโดยกฎชิลลิ่งทั้ง 3 แบบจากการตอบผิดตำแหน่งที่หนึ่งมีค่าแตกต่างกันน้อยมาก เพราะเงื่อนไขการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องกับเงื่อนไขการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้งทำให้การตอบผิดตำแหน่งที่หนึ่งเป็นตำแหน่งเดียวกัน จึงมีค่าความเชื่อมั่นเท่ากันเท่ากับ .96 และเงื่อนไขการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัวจากการตอบผิดตำแหน่งที่หนึ่ง ก็คือตำแหน่งเดียวกันกับการตอบผิดตำแหน่งที่สองของการตรวจให้คะแนนโดยกฎชิลลิ่งแบบคะแนนศูนย์ต่อเนื่อง จึงมีค่าความเชื่อมั่นเท่ากันเท่ากับ .95 และจากการทดสอบเป็นรายคู่ จากการตอบผิดตำแหน่งที่หนึ่งกับการตอบผิดตำแหน่งที่สอง

ของการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องในข้อที่ 1 มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติแสดงให้เห็นว่าค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งของวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบ คือ .96, .96 และ .95 มีค่าไม่แตกต่างกัน และ 2) เป็นข้อเท็จจริงที่ว่า การทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีของเซฟเฟ (Scheffe's Test) เป็นวิธีการที่มีลักษณะเข้มงวดมากกว่าวิธีอื่น กล่าวคือ จะมีเกณฑ์ในการเปรียบเทียบเป็นรายคู่สูงกว่าวิธีการอื่น ๆ ดังนั้น อาจจะพบว่าไม่มีคู่ใดแตกต่างกันเลย ถึงแม้ว่า F - Distribution ในการทดสอบรวมมีความแตกต่างกัน (Ferguson G. A. & Yoshio Takane. 1989 : 334)

5. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบ จากการตอบผิดตำแหน่งที่สองของแบบทดสอบ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว และการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว สอดคล้องกับสมมติฐานที่ตั้งไว้ ส่วนการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เพราะค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งกับตำแหน่งที่สอง ของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องและกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง มีค่าไม่แตกต่างกัน และค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่หนึ่งของแบบทดสอบที่ตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องและกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง มีค่าเท่ากัน

6. ค่าความเชื่อมั่นระหว่างวิธีการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 3 แบบ ของแบบทดสอบจากการตอบผิดตำแหน่งที่สามมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ การตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง และการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องเพียงตัวเดียวหลาย ๆ ครั้ง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ส่วนการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่อง กับการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้ อาจเนื่องมาจากยิ่งทำข้อสอบมาก ข้อความยากก็ยิ่งเพิ่มมากขึ้น และโอกาสการตอบผิดที่มีมากขึ้นตามเงื่อนไขการตรวจที่เปิดโอกาสให้ตอบผิดติดต่อกันได้มากขึ้นถึง 3 ครั้ง คือ (000) และ (00, 00, 00) จึงทำให้ค่าความเชื่อมั่นจากการตอบผิดตำแหน่งที่สามของการตรวจให้คะแนนโดยกฎซิลลิงทั้ง 2 แบบ ใกล้เคียงกัน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรศึกษาเกี่ยวกับวิธีการตรวจให้คะแนนโดยกฎซิลลิงที่แตกต่างกันทั้ง 3 แบบกับนักเรียนระดับประถมศึกษาชั้นอื่น ๆ ว่าจะมีผลสอดคล้องกันหรือไม่กับการศึกษาวิจัยครั้งนี้
2. ควรทำการวิจัยในทำนองนี้โดยเจาะจงในกลุ่มตัวอย่างที่มีความสามารถต่างกันโดยจำแนกออกเป็นนักเรียนกลุ่มเก่ง นักเรียนกลุ่มปานกลาง และนักเรียนกลุ่มอ่อน
3. ควรทำการวิจัยที่ละพฤติกรรมเพื่อนำผลมาเปรียบเทียบว่าเหมือนหรือต่างกัน

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ในการสอบแข่งขัน หรือสอบคัดเลือกกิจกรรมต่าง ๆ ของโรงเรียน หน่วยงานควรสร้างแบบทดสอบที่มีคุณภาพโดยจัดเรียงข้อคำถามตามค่าความยากจากข้อง่ายไปยากและใช้วิธีการตรวจให้คะแนนโดยกฎซิลลิงเพื่อหาความสามารถที่แท้จริงของผู้สอบ
2. จากผลการวิจัยครั้งนี้พบว่า ควรใช้วิธีการตรวจให้คะแนนโดยกฎซิลลิงแบบคะแนนศูนย์ต่อเนื่องมากกว่า 1 ตัว เพราะมีความแตกต่างกันในทุกตำแหน่งการตอบผิด และให้โอกาสนักเรียนในการตอบผิดได้มากกว่า
3. สำหรับนักเรียนระดับประถมศึกษาถ้าแบบทดสอบที่ใช้มีจำนวนข้อสอบมากข้อควรแบ่งข้อสอบออกเป็นตอนหรือเป็นฉบับหลาย ๆ ฉบับและเว้นระยะเวลาในการทดสอบ

บรรณานุกรม

บรรณานุกรม

- เกริกชัย ฮวบเจริญ. (2525). การเปรียบเทียบการตอบ และการตรวจให้คะแนนแบบใหม่ และวิธีทดสอบความมั่นใจ และแบบธรรมดาที่มีต่อค่าความเชื่อมั่น และคะแนนการเดาของแบบทดสอบเลือกตอบ. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). 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. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย ศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ชวาล แพร่สกุล. (2516). เทคนิคการวัดผล. พิมพ์ครั้งที่ 5. วัฒนาพานิช.
- _____. (2517). การทดสอบเพื่อค้น และพัฒนาสมรรถภาพ. หน้า 46 - 47. กรุงเทพฯ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2518). เทคนิคการวัดผล. พิมพ์ครั้งที่ 6. กรุงเทพฯ ฯ : วัฒนาพานิช.
- เดือนใจ เศรษฐลักโก. (2511). การศึกษาถึงการจัดระเบียบข้อสอบด้วยวิธีต่าง ๆ ที่จะส่งผลต่อความสามารถของนักเรียนชั้น ป.7. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ทวี ทองคำ. (2526). การเปรียบเทียบค่าความเที่ยง ความตรง และอำนาจจำแนกของแบบทดสอบชนิดเลือกตอบที่ใช้คำสั่งและวิธีการให้คะแนนที่ต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ธีรศักดิ์ อินทรมาตย์. (2520). อิทธิพลของวิธีการตอบแบบทดสอบปรนัยชนิดเลือกตอบวิธีต่าง ๆ ที่มีต่อค่าความเชื่อมั่น ความเที่ยงตรงและปริมาณการเดา. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.

- นภา กาญจนกิจโสภณ. (2528). การเปรียบเทียบคะแนนสอบระหว่างการให้น้ำหนักคะแนนรายข้อเท่ากันตามทฤษฎีมาตรฐานเดิมกับการให้น้ำหนักคะแนนรายข้อต่างกันตาม ทฤษฎีการตอบสนองข้อคำถาม. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์วิทยาลัย. ถ่ายเอกสาร.
- นิโลบล นิ่มกิ่งรัตน์. (2519). อิทธิพลของการจัดเรียงลำดับข้อคำถามที่จำแนกตามสมรรถภาพสมองที่มีผลต่อคุณภาพของแบบทดสอบคณิตศาสตร์ และภาษาไทย ม. 3. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- แน่นน้อย เพชรรัตน์. (2521). ผลการเรียงลำดับข้อคำถามด้วยวิธีต่างกันที่มีผลต่อคุณภาพของแบบทดสอบวิชาทักษะกระบวนการวิทยาศาสตร์. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- บังอร กมลวัฒนา. (2542). การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน. ปรินูญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- บุญชม ศรีสะอาด. (2540). การวิจัยทางการวัดผลและประเมินผล. กรุงเทพฯ ฯ : สุวีริยาสาส์น
- บุญเชิด ภิญโญอนันตพงษ์. (2521). การวัดและประเมินผลการศึกษา : ทฤษฎีและการประยุกต์. คณะศึกษาศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2526). การสอนแบบอิงเกณฑ์ : แนวคิดและวิธีการ. กรุงเทพฯ ฯ : ภาควิชาพื้นฐานการศึกษา. คณะศึกษาศาสตร์. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- _____. (2533). "Congeneric Part Reliability," การวัดผลการศึกษา. 12 (34) : 28 - 31; พฤษภาคม - สิงหาคม.
- _____. (2536). "Congeneric Model," การวัดผลการศึกษา. 11(42) : 8 - 20 ; มกราคม - เมษายน.
- _____. (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์. ปรินูญานิพนธ์ กศ.ด. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- _____. (2538). "แบบทดสอบคะแนนจริงสัมพันธ์:สัมประสิทธิ์ r_B ," วารสารการวัดผลการศึกษา. กรุงเทพฯ ฯ.

- บุญเชิด ภิญโญอนันตพงษ์. (2542). สัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน. กรุงเทพฯ ฯ : ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- พนิชา สวัสดิ์มงคล. (2530). การเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบในการวัดระดับความรู้ที่ต่างกันวิธีโดยให้คะแนนต่างกันในกลุ่มผู้สอบที่มีผลสัมฤทธิ์ทางการเรียนต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พินิจ อุไรรักษ์. (2533). ผลของวิธีการให้คะแนนที่มีต่อคะแนนสอบ ค่าความเชื่อมั่น และค่าความเที่ยงตรง. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- เพ็ญศรี สว่างเนตร. (2520). ความเที่ยงตรงของแบบสอบสองชนิดเลือกตอบด้วยเทคนิคการให้คะแนนที่ต่างกัน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- เพลินพิศ คุณคำ. (2532). การประมาณความเชื่อมั่นและปริมาณการเดาของแบบทดสอบปรนัยเลือกตอบที่มีวิธีการตอบและการตรวจให้คะแนนวิธี 0 - 1, วิธีของคูมปัส และ วิธีของอนันต์. ปรินญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- พจน์ สะเพียรชัย. (2515). " ทฤษฎีการตรวจให้คะแนนกับการเดา," พัฒนาวัดผล 8. หน้า 21 - 25. กรุงเทพฯ ฯ : อักษรสัมพันธ์.
- มหาวิทยาลัย, ทบวง. (2524). ชุดการเรียนการสอนสำหรับครูคณิตศาสตร์. 335 หน้า.
- มานะ คิ้วเจริญ. (2539). การศึกษาผลการตรวจให้คะแนนโดยกฎซึลลิ่งที่กำหนดจุดซึลลิ่งต่างกันของแบบทดสอบคณิตศาสตร์ ที่มีการจัดเรียงข้อคำถามตามค่าความยาก และ พฤติกรรมการเรียนรู้. ปรินญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- เยาวลักษณ์ น้อยนาแสง. (2535). การเปรียบเทียบคุณภาพของข้อสอบความถนัดทางการเรียนแบบปรนัยชนิดเลือกตอบซึ่งเป็นผลจากวิธีการตอบและการตรวจให้คะแนนที่แตกต่างกัน. ปรินญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- รัชณี กันภัย. (2543). ผลของการจัดเรียงข้อสอบต่างกันและกำหนดเวลาให้ทำต่างกันที่มีต่อค่าความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบที่สอบด้วยไมโครคอมพิวเตอร์. ปรินญานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.

- ล้วน สายยศ . (2517). “การวัดผลสัมฤทธิ์วิชาคณิตศาสตร์,” เอกสารการประชุมวิชาการครั้งที่ 2. หน้า 36 - 39. กรุงเทพฯ ฯ : โรงพิมพ์ศาสนา.
- _____. (2526). “วิเคราะห์ผลการเดาข้อสอบก่อนสอบและหลังสอบกับความสามารถจริง,” วารสารการวัดผลการศึกษา. กรุงเทพฯ ฯ.
- ล้วน สายยศและอังคณา สายยศ. (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ ฯ : ชมรมเด็ก.
- _____. (2540). สถิติวิทยาทางการวิจัย. พิมพ์ครั้งที่ 3. กรุงเทพฯ ฯ : สุวีริยาสาส์น.
- วัชรารภรณ์ จิตรมาศ. (2538). การเปรียบเทียบคุณภาพแบบทดสอบเลือกตอบชนิดถูกผิดหลายตัวเลือกที่มีจำนวนตัวเลือกและวิธีตรวจให้คะแนนต่างกัน. ปรินูญานินพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศศิธร สุวรรณสุข. (2519). การศึกษาค่าความเที่ยงตรง ค่าความเชื่อมั่นของแบบทดสอบที่เรียงลำดับคำถามต่าง ๆ กันในวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 7. ปรินูญานินพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศิริวิมล สุขสมบูรณ์. (2541). ผลการวิเคราะห์แบบแผนการตอบข้อสอบของแบบทดสอบเลือกตอบที่มีรูปแบบการจัดเรียงข้อความต่างกัน. ปรินูญานินพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ศึกษาธิการ, กระทรวง. (2533 ก). คู่มือหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ ฯ : โรงพิมพ์การศาสนา.
- _____. (2533 ข). หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ ฯ : โรงพิมพ์การศาสนา.
- _____. (2534). คู่มือการประเมินผลการเรียนตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ ฯ : โรงพิมพ์คุรุสภาลาดพร้าว
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2534). คู่มือครูคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6. หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กระทรวงศึกษาธิการ : โรงพิมพ์คุรุสภา.
- สว่างค์ ร่มเงิน. (2541). การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ที่มีรูปแบบการจัดเรียงข้อความตามค่าความยากทางสถิติและพฤติกรรมการเรียนรู้ต่างกัน. ปรินูญานินพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- สิริพร พุ่มแก้ว. (2521). ผลของการเรียงลำดับข้อสอบต่อค่าความเที่ยงตรงของแบบทดสอบและผลสัมฤทธิ์ทางการเรียน. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

- สุพรรณ สุกมลสันต์. (2528). “ผลการกำหนดน้ำหนักคะแนนต่างคุณภาพของแบบสอบ,” วารสาร
วิธีวิทยาการวิจัย. กรุงเทพฯ ฯ : จุฬาลงกรณ์มหาวิทยาลัย, พฤษภาคม - สิงหาคม.
- สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ. (2538). การวัดและประเมินผลในชั้นเรียนกลุ่ม
ทักษะคณิตศาสตร์. เอกสารกองวิชาการ ลำดับที่ 18 / 2538. กรุงเทพฯ ฯ : โรงพิมพ์คุรุ
สภาลาดพร้าว.
- สำนักงานคณะกรรมการการศึกษาแห่งชาติ. สำนักนายกรัฐมนตรี. (2540). แผนพัฒนาการศึกษา
แห่งชาติฉบับที่ 8. (พ.ศ. 2540 - 2544). กรุงเทพฯ ฯ : อรรถพลการพิมพ์.
- สำราญ มีแจ้ง. (2525). ผลของคำสั่งการใช้คะแนนที่ต่างกันต่อค่าความเที่ยง ความตรง และ
อำนาจจำแนกของแบบทดสอบชนิดเลือกตอบ. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการ
ศึกษา). กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- สำเร็จ บุญเรืองรัตน์. (2525). “การพัฒนาทฤษฎีลาเท้นเพื่อวิเคราะห์ข้อสอบ,” วารสารการวัดผล
การศึกษา 4 (2) : 1 – 13 ; กันยายน – สิงหาคม.
- สำเร็จ บุญเรืองรัตน์. (2529). ทฤษฎีการวัดและการประเมินผลการศึกษา. พิมพ์ครั้งที่ 2
กรุงเทพฯ ฯ : สยามศึกษา.
- สมพงษ์ อัสวศุภฤกษ์. (2542). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจและ
ความยาวของแบบทดสอบต่างกัน. วิทยานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา).
กรุงเทพฯ ฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- อนันต์ ศรีโสภ. (2520) . การวัดผลและประเมินผล. กรุงเทพฯ ฯ : ไทยวัฒนาพานิช.
- _____. (2524). การวัดผลและการประเมินผลการศึกษา. พิมพ์ครั้งที่ 2. กรุงเทพฯ ฯ : ไทยวัฒนา
พานิช.
- _____. (2526). “เราจะพัฒนาวิธีการตอบและตรวจให้คะแนนข้อสอบแบบเลือกตอบได้อย่างไร,”
สามัญศึกษา. 13 - 19 : เมษายน.
- _____. (2532). การทดลองรูปแบบใหม่ของข้อสอบแบบเลือกตอบ. กรุงเทพฯ ฯ : คณะศึกษ
ศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- อัมพร ยิ้มยิ้ม. (2533). ผลของการตรวจให้คะแนน โดยกฎซัลลิ่งที่กำหนดจุดซัลลิ่งต่างกันที่มีต่อ
ค่าการเดา และลักษณะโค้งอินเฟอร์เมชันของแบบทดสอบวัดผลสัมฤทธิ์วิชา
คณิตศาสตร์. วิทยานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ : บัณฑิต
วิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- อรรณพ ดัชนีเจริญรัตน์ . (2517). การศึกษาวิธีการตอบ และการตรวจให้คะแนนแก่ข้อสอบ
ปรนัยแบบเลือกตอบ. วิทยานิพนธ์ กศ.ม. (วิชาเอกการวัดผลการศึกษา). กรุงเทพฯ ฯ
: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.

- อวยพร วิบูลย์กาญจน์. (2526). *การเปรียบเทียบผลการวิเคราะห์แบบสอบอุปมาอุปไมยด้วยคลาสสิกอลโมเดลกับราชคโมเดล*. วิทยานิพนธ์ ค.ม. (ภาควิชาวิจัยการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- อุทุมพร จามรมาน. (2536). *ข้อสอบ : การสร้างและการพัฒนา*. กรุงเทพฯ : ฟันนี่พับลิชชิง.
- Abu - Sayf, F.X. and James J. Diamond. (1976). "Effect of Confidence Level in Multiple - Choice Test Answer on Reliability and Validity of Score," *Journal of Educational Resarch*. 70 : 62 ; Novenber-December.
- Albert, Wilfred George. (1970). "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 ," *Dessertation Abstracts International*. 31 (4) : 1619 A ; October.
- 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.
- Coombs Clydes H, J.E. Milholand and F.B. womer. (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 J. Schmid. (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.
- 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.

- Flanger, R. Melton, R., and Myers, C. (1968). "Item Rearrangement under Typical Test Conditional," *Educational and Psychological Measurement*. London : Sage Publications, Inc. 28, p. 813 - 824.
- Garret, Henry E. and R.S. Wood worth. (1967). *Statistical in Psychology and Education*. David Mckay Company, Inc.
- Gibbon, Jean. D., Olkin, Inqram, and Sobel, Melton. (1979). "A Subtest Selection technique for scoring Item on a Multiple Choice test," *Psychometrika* . 44 : 259 - 270 ; September.
- 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.
- Klosner, Naomi Certner & Gellman, Estelle Klitnick. (1973). "The Effect of Item Arrangment on classroom Test Performance : Implication for Content Validity," *Educational and Psychological Measurement*. London : Sage Publications, Inc. 33 : 413 - 418.
- Lane, S., J.R , David & Bull, Kay S. (1987). "The Effects of Knowledge of Item Arrangment, Gender, and Statistical and Cognitive Item Diffculty an Test Performance," *Educational and Psychological Measurement*. London : Sage Publications, Inc. 865 - 879.
- Lindqusit, 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 J. Lehmann. (1984). *Measurement and Evaluation in Educational and Psychology*. 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 J. Brunza. (1975). " Effect of Confidence Weighted Scoring System on Measures of Test Reliability and Validity," *Educational and Psychological Measurement*. 35 : 73 - 78 ; Spring.
- Rasch, G. (1960.) *Probabilistic models for some Intelligence and Attainment Tests*. Copenhagen, Denmark : Denmark's Paedagogiske Institute.
- Scheffe', H. (1959). *The Analysis of Variance*. New York : John Wiley and Sons. Inc.
- Stanley, Julian C. and Kenneth D. Hopkins. (1972). *Educational and Psychology Evaluation*. (5 th. Ed). Englewoods Cliff ; Prentice - 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.
- 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

ดร. สุวพร เข้มเฮง	ภาควิชาการวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร. สมบัติ สุวรรณพิทักษ์	ผู้อำนวยการกองพัฒนา กรรมการศึกษานอกโรงเรียน กระทรวงศึกษาธิการ
อาจารย์ชวลิต รวยอาจัน	ภาควิชาการวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์ขรรค์ชัย คงเสน่ห์	นักวิชาการศึกษา 8 กรรมการศึกษานอกโรงเรียน กระทรวงศึกษาธิการ
อาจารย์ธานี เขียนมัน	อาจารย์ 2 ระดับ 7 โรงเรียนบ้านทุ่งประทุน ตำบลท่าเรือ อำเภอท่ามะกา จังหวัดกาญจนบุรี

ภาคผนวก ข

ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

(Index of Consistency : IOC)

ตาราง ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับพฤติกรรมการเรียนรู้

ข้อ	IOC	ข้อ	IOC	ข้อ	IOC	ข้อ	IOC	ข้อ	IOC	ข้อ	IOC
1	1	29 *	1	57	0.8	85	1	113	1	141 *	1
2 *	1	30	1	58	1	86	1	114 *	1	142	1
3	1	31 *	1	59 *	1	87	1	115 *	1	143 *	1
4	0.8	32	1	60	1	88	1	116 *	1	144	1
5	1	33	1	61	1	89	1	117 *	1	145	1
6	1	34 *	1	62 *	1	90	1	118 *	1	146	1
7 *	1	35 *	1	63	1	91	1	119	1	147	1
8 *	1	36	1	64 *	1	92	1	120	1	148	1
9 *	1	37	1	65	1	93	1	121	1	149	1
10	1	38 *	1	66	1	94	1	122	1	150 *	1
11 *	1	39	1	67	1	95	1	123	1	151 *	1
12	1	40	1	68 *	1	96 *	1	124	1	152	1
13	0.6	41 *	1	69	1	97 *	1	125	1	153 *	1
14	1	42	1	70 *	1	98 *	1	126 *	1	154 *	1
15	1	43 *	1	71 *	1	99	1	127 *	1	155	1
16	1	44	1	72 *	1	100 *	1	128	1	156	1
17 *	1	45	1	73	1	101 *	1	129	1	157	1
18 *	1	46 *	1	74	0.8	102	1	130	1	158 *	1
19 *	1	47	1	75 *	0.8	103	1	131	1	159	1
20	1	48 *	1	76	0.8	104	1	132 *	1	160 *	1
21	1	49	1	77 *	1	105 *	1	133 *	1	161 *	1
22	1	50 *	1	78	1	106 *	1	134	1	162 *	1
23	1	51	1	79 *	1	107	1	135 *	1	163	1
24	1	52	1	80	1	108	1	136	1	164 *	1
25 *	1	53	1	81	1	109	1	137	1		
26	1	54 *	0.8	82	1	110	1	138 *	1		
27	1	55	0.8	83	1	111	1	139	1		
28	1	56	0.8	84	1	112	1	140 *	1		

* หมายถึง ข้อที่คัดเลือกไว้ใช้ในการวิจัยโดยพิจารณาให้ครอบคลุมตามพฤติกรรมการเรียนรู้ของวิลสัน

ภาคผนวก ค

**ค่าความยาก (b) ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 ที่วิเคราะห์โดยราชคโมเดล (Rasch Model)
ที่เรียงจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด จำนวน 50 ข้อ**

ตาราง ค่าความยาก (Item Difficulty : b) ของแบบทดสอบวัดวัดผลสัมฤทธิ์ทางการเรียนวิชา
คณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่วิเคราะห์ตามแบบราชคโมเดล (Rasch Model) จาก
การทดลองใช้ (Try out)

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1	-4.565	21	-0.228	41	-0.977
2	-1.082	22	1.937	42	-1.706
3	1.409	23	-2.285	43	1.144
4	-0.125	24	-0.114	44	0.640
5	0.399	25	0.143	45	-0.561
6	-1.557	26	-1.776	46	0.920
7	-0.036	27	0.231	47	1.347
8	0.675	28	-1.420	48	0.711
9	0.535	29	0.365	49	0.723
10	-0.297	30	-0.415	50	-0.919
11	1.608	31	-0.002	51	-0.475
12	0.997	32	-0.274	52	-0.947
13	0.287	33	-0.309	53	-0.573
14	-0.714	34	0.894	54	0.451
15	-0.171	35	1.347	55	0.054
16	0.320	36	1.076	56	0.882
17	-0.807	37	-1.129	57	0.932
18	0.087	38	0.807	58	-0.451
19	1.457	39	-1.978	59	1.813
20	1.036	40	0.971	60	0.593

ตาราง ค่าความยาก (Item Difficulty : b) ของแบบทดสอบวัดวัดผลสัมฤทธิ์ทางการเรียนวิชา
คณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ใช้ในการวิจัย

ข้อที่	ค่าความยาก (b)	ข้อที่	ค่าความยาก (b)
1	-1.978	26	0.143
2	-1.776	27	0.231
3	-1.557	28	0.320
4	-1.129	29	0.365
5	-1.082	30	0.399
6	-0.977	31	0.535
7	-0.947	32	0.593
8	-0.919	33	0.640
9	-0.807	34	0.675
10	-0.714	35	0.711
11	-0.573	36	0.723
12	-0.561	37	0.882
13	-0.475	38	0.894
14	-0.451	39	0.920
15	-0.415	40	0.932
16	-0.309	41	0.971
17	-0.297	42	0.997
18	-0.274	43	1.036
19	-0.228	44	1.076
20	-0.171	45	1.144
21	-0.125	46	1.347
22	-0.114	47	1.409
23	-0.002	48	1.457
24	0.054	49	1.608
25	0.087	50	1.937

ภาคผนวก ง
แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6

คำชี้แจง

1. แบบทดสอบมีทั้งหมด 1 ฉบับ จำนวน 50 ข้อ ใช้เวลา 80 นาที
2. คำถามทุกข้อเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุดเพียงข้อเดียว จากข้อ ก ข ค หรือ ง โดยกาเครื่องหมาย $\times$ ลงในช่องสี่เหลี่ยมใต้ข้อที่นักเรียนเลือกในกระดาษคำตอบ ดังตัวอย่างต่อไปนี้

(0) ข้อใดเป็นจริง

ก. $2 + 3 = 3 + 2$

ข. $5 \times 0 = 2 + 3$

ค. $(2 + 3) + 4 = 2 + (3 + 5)$

ง. $10 - 8 - 2 = 10 - (8 - 2)$

จะเห็นว่าคำตอบที่ถูกต้องคือ ข้อ ก ให้นักเรียนกาเครื่องหมาย $\times$ ลงในช่องสี่เหลี่ยมให้ตรงกับข้อ (0) ได้ตัวเลือก ก ในกระดาษคำตอบดังนี้

ข้อ	ก	ข	ค	ง
0	$\times$			

3. ถ้านักเรียนต้องการเปลี่ยนคำตอบใหม่ให้ชัดเจน (=) ทับรอบเดิมก่อนแล้วเลือกคำตอบใหม่ เช่น ต้องการเปลี่ยนคำตอบจากข้อ ก เป็นข้อ ง ให้ทำดังนี้

ข้อ	ก	ข	ค	ง
0	~~$\times$~~			$\times$

4. ให้นักเรียนทศเลขในกระดาษที่แจกให้
5. ห้ามนักเรียนทำเครื่องหมายใด ๆ ลงในแบบทดสอบนี้
6. ให้นักเรียนส่งแบบทดสอบและกระดาษคำตอบคืนกรรมการคุมสอบ
7. ให้นักเรียนทำข้อสอบทุกข้อ เพื่อประโยชน์ของตัวนักเรียนเอง

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6

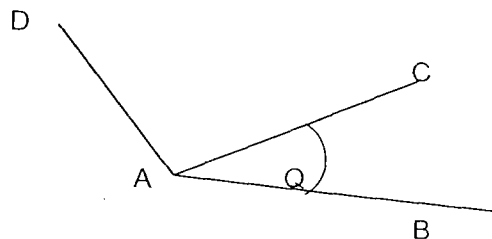
1. แก้วมีเงิน 515 บาท ก้อยมีเงิน 1,011 บาท ไก่มีเงิน 1,500 บาท
แก้ว ไก่ ก้อย มีเงินรวมกันประมาณกี่บาท

ก. 3,000 บาท
ข. 3,150 บาท
ค. 3,500 บาท
ง. 4,000 บาท

2. ข้อใดคือสมการที่เป็นจริง

ก. $17 + 19 = 26$
ข. $100 - 20 = 80$
ค. $125 \div 5 = 22$
ง. $73 \times 20 = 1,430$

3. จากรูปที่กำหนดให้ ข้อใดเรียกชื่อมุม Q ได้ถูกต้อง



ก. $\widehat{CAB}$
ข. $\widehat{DAC}$
ค. $\widehat{DAB}$
ง. $\widehat{CAD}$

4. ผงชูรส 3 ห่อ ๆ แรกหนัก 10.255 กรัม ห่อที่สองหนัก 10.712 กรัม ห่อที่สามหนัก 10.521 กรัม จงจัดลำดับห่อผงชูรสจากที่มีน้ำหนักมากไปยังห่อที่มีน้ำหนักน้อยที่สุด

ก. 10.712, 10.521, 10.255
ข. 10.712, 10.255, 10.521
ค. 10.255, 10.521, 10.721
ง. 10.521, 10.255, 10.721

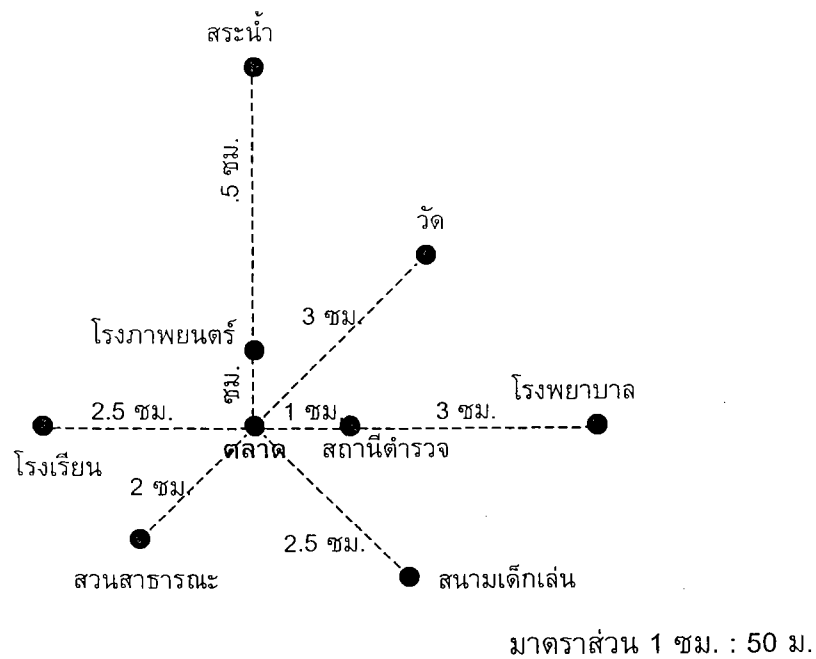
5. ข้อใดเป็นสมการ

- ก. $5 + 10 < 50$
- ข. $42 \times 7 > 12$
- ค. $3 + 14 \neq 1 + 43$
- ง. $10 - (9 - 7) = (10 - 9) - 7$

6. ผลบวกของจำนวนเลขคู่ตั้งแต่ 1 – 10 กับผลบวกของจำนวนเลขคี่ตั้งแต่ 1 – 10 มีค่าต่างกันอยู่เท่าไร

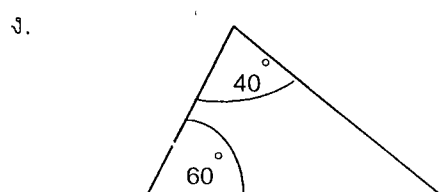
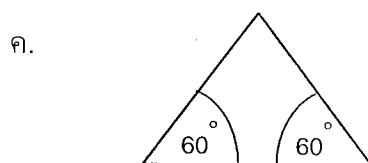
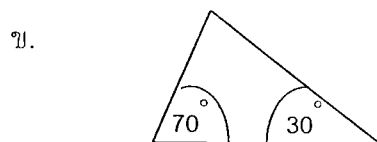
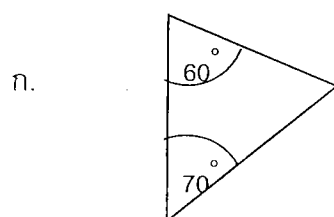
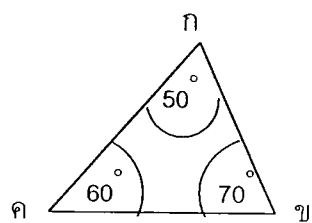
- ก. 0
- ข. 1
- ค. 5
- ง. 10

7. จากแผนผังที่กำหนดให้ สถานที่ใดอยู่ห่างจากตลาดเป็นระยะทาง 200 เมตร



- ก. วัด
- ข. โรงเรียน
- ค. โรงพยาบาล
- ง. สโมสรเด็กเล่น

8. จากรูปที่กำหนดให้ รูปในข้อใดมีขนาดมุมเท่ากับ คักข



9. ถ้ากำหนดเงินรางวัลให้แก่ผู้แข่งขัน

รางวัลที่ 1 เป็นเงิน 25,000 บาท

รางวัลที่ 2 เป็นเงิน 10,080 บาท

รางวัลที่ 3 เป็นเงิน 7,750 บาท

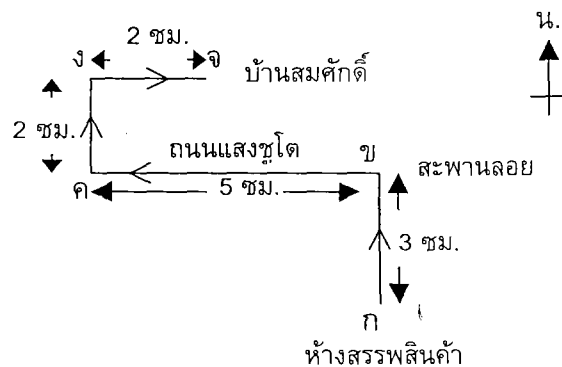
แต่ปรากฏว่ามีผู้ได้รับรางวัลที่ 2 ถึง 2 ราย ส่วนรางวัลที่ 3 ไม่มีผู้ได้รับ จะต้องหาเงินรางวัลเพิ่มขึ้นอีกกี่บาท

- ก. 2,330 บาท
- ข. 19,780 บาท
- ค. 27,695 บาท
- ง. 30,025 บาท

10. 77.051 เขียนในรูปการกระจายได้ดังข้อใด

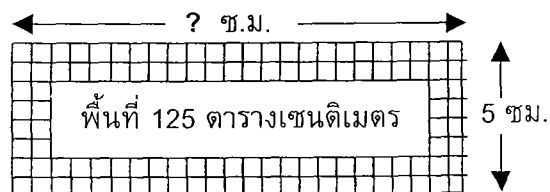
- ก. $70 + 7 + 0.5 + 0.001$
- ข. $70 + 7 + 0.05 + 0.001$
- ค. $70 + 0.7 + 0.05 + 0.01$
- ง. $70 + 0.7 + 0.05 + 0.001$

11. จากแผนผังที่กำหนดให้ สมศักดิ์ต้องเดินทางจากห้างสรรพสินค้ากลับบ้าน เป็นระยะทางเท่าไร เมื่อใช้มาตราส่วน 1 ซม. : 250 เมตร



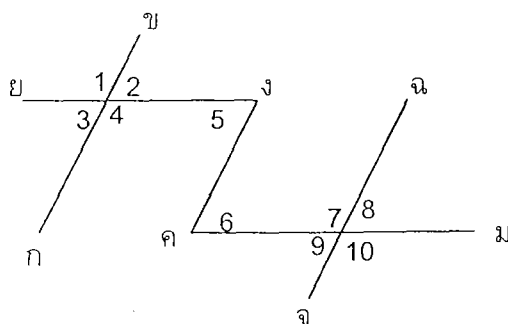
- ก. 3 กิโลเมตร
- ข. 2 กิโลเมตร
- ค. 2 กิโลเมตร 250 เมตร
- ง. 1 กิโลเมตร 750 เมตร

12. จากรูปที่กำหนดให้ รูปสี่เหลี่ยมมีด้านยาวกี่เซนติเมตร



- ก. 20 เซนติเมตร
- ข. 24 เซนติเมตร
- ค. 25 เซนติเมตร
- ง. 26 เซนติเมตร

13. จากรูป กำหนดให้ $\overline{กข} \parallel \overline{คง} \parallel \overline{จณ}$ และ $\overline{ยง} \parallel \overline{คม}$ มี $\angle 2 = 50^\circ$ มุมในข้อใด มีขนาดมุมเท่ากับ 50° ทุกมุม



- ก. $\angle 1, \angle 5, \angle 9$
- ข. $\angle 3, \angle 5, \angle 9$
- ค. $\angle 4, \angle 5, \angle 9$
- ง. $\angle 6, \angle 8, \angle 10$

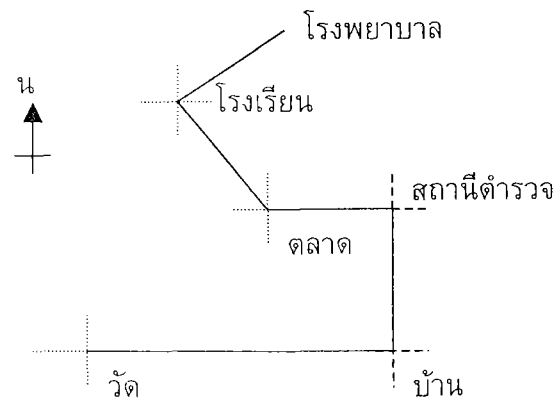
14. เส้นผมของวิกานดายาว 1.39 เซนติเมตรต่อเดือน ในเวลา 1 ปี ถ้าวิกานดาไม่ตัดผม เส้นผมของวิกานดาจะยาวเท่าไร

- ก. 13.39 เซนติเมตร
- ข. 16.68 เซนติเมตร
- ค. 31.39 เซนติเมตร
- ง. 41.70 เซนติเมตร

15. ข้อใดเป็นการแยกตัวประกอบ

- ก. 2×10^3
- ข. $2 \times 8 \times 9$
- ค. $34 \times 10 \times 10$
- ง. $3 \times 5 \times 7 \times 13$

16. จากรูปที่กำหนดให้ บ้านอยู่ทางทิศใดของโรงเรียน

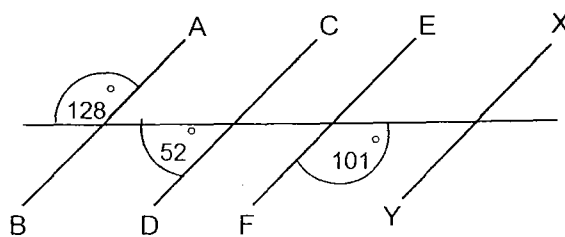


- ก. ทิศตะวันตกเฉียงใต้
- ข. ทิศตะวันออกเฉียงใต้
- ค. ทิศตะวันตกเฉียงเหนือ
- ง. ทิศตะวันออกเฉียงเหนือ

17. เศษส่วนข้อใดมีค่าเท่ากับ $\frac{8}{12}$

- ก. $\frac{2}{3}, \frac{4}{6}, \frac{16}{24}$
- ข. $\frac{1}{4}, \frac{4}{6}, \frac{16}{12}$
- ค. $\frac{4}{12}, \frac{4}{3}, \frac{8}{24}$
- ง. $\frac{8}{24}, \frac{2}{3}, \frac{24}{36}$

18. จากรูปที่กำหนดให้ ข้อใดสรุปได้ถูกต้อง



ก. $\overline{AB} // \overline{CD}$

ข. $\overline{AB} // \overline{EF}$

ค. $\overline{AB} // \overline{XY}$

ง. $\overline{CD} // \overline{EF}$

19. ต้นมะม่วงสูงกว่าต้นมะพร้าว 2.50 เมตร ต้นมะพร้าวเตี้ยกว่าต้นมะยม 1.25 เมตร
ต้นมะพร้าวสูง 0.5 เมตร ต้นมะยมสูงกี่เมตร

ก. 1.75 เมตร

ข. 2.00 เมตร

ค. 2.25 เมตร

ง. 2.50 เมตร

20. ข้อใดต่อไปนี้ถูกต้อง

ก. $1.7230 = \frac{1723}{10000}$

ข. $1.7230 = 1 \frac{7230}{100}$

ค. $1.7230 = 1 \frac{7230}{1000}$

ง. $1.7230 = 1 \frac{7230}{10000}$

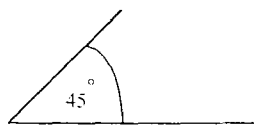
21. จำนวนในข้อใดเป็นจำนวนเฉพาะทุกจำนวน

- ก. 1, 2, 3
- ข. 2, 3, 5
- ค. 5, 7, 9
- ง. 6, 7, 11

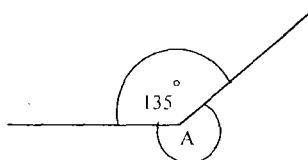
22. ค่าประมาณใกล้เคียงจำนวนเต็มร้อยของ 639,905 คือข้อใด

- ก. 630,000
- ข. 639,000
- ค. 639,900
- ง. 640,000

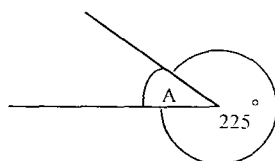
23. มุม A ในข้อใด มีขนาดเท่ากับมุมที่กำหนดให้



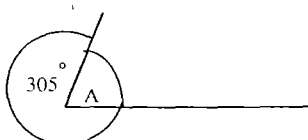
ก.



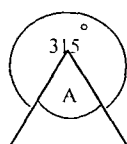
ข.



ค.



ง.



24. รั้วบั้งเส้นหนึ่งยาว $100\frac{1}{2}$ เมตร ตัดปลาย $4\frac{1}{2}$ เมตร ที่เหลือแบ่งเป็น 12 ท่อน
ท่อนละเท่าๆ กัน จะได้รับบั้งยาวท่อนละเท่าไร

- ก. 4 เมตร
ข. 8 เมตร
ค. 84 เมตร
ง. $8\frac{3}{4}$ เมตร

25. ให้ x, y และ z เป็นจำนวนนับ ซึ่ง x เป็น 2 เท่าของ z , y เป็น 3 เท่าของ z
ถ้า $y = 6$ จะได้ $x + y + z$ มีค่าเท่าไร

- ก. 12
ข. 14
ค. 16
ง. 22

26. ซื้อเสื้อมาราคาตัวละ 320 บาท ซื้อกางเกงราคาตัวละ 950 บาท ให้ธนบัตรราคา
ใบละ 500 บาท จำนวน 3 ใบจะได้รับเงินทอนกี่บาท จงเขียนเป็นประโยคสัญลักษณ์

- ก. $(950 - 320) - 500 \times 3 = \square$
ข. $(950 + 320) - 500 \times 3 = \square$
ค. $500 \times 3 - (950 + 320) = \square$
ง. $500 \times 3 - (950 - 320) = \square$

27. จากสมการ $\frac{m}{3} = 10$ ข้อใดเป็นวิธีทำให้สมการเป็นจริง

- ก. $\frac{m}{3 \times 3} = 10 \times 3$
ข. $\frac{m}{3} \times 3 = 10 \times 3$
ค. $\frac{m}{3 \times 3} = \frac{10}{3}$
ง. $\frac{m}{3} - 3 = 10 - 3$

28. $(100.5 \times 2.5) \times 0.1 = \square$

- ก. 2.5125
- ข. 25.125
- ค. 251.25
- ง. 2,512.5

29. ตัวประกอบเฉพาะทั้งหมดของ 60 คือข้อใด

- ก. 1, 3, 5
- ข. 2, 3, 5
- ค. 2, 6, 5
- ง. 2, 5, 10

30. เมื่อ ก เป็นจำนวนนับ ข และ ค เป็นจำนวนเฉพาะ ข้อใดเป็นการแยกตัวประกอบ

- ก. $ก = ข + ค$
- ข. $ก = ข - ค$
- ค. $ก = ข \times ค$
- ง. $ก = ข \div ค$

31. ทิศที่ทำมุม 135 องศา กับทิศตะวันตกเฉียงใต้คือทิศในข้อใด

- ก. ตะวันออกกับเหนือ
- ข. ตะวันออกกับตะวันตก
- ค. ตะวันออกเฉียงเหนือกับตะวันตก
- ง. ตะวันออกกับตะวันออกเฉียงเหนือ

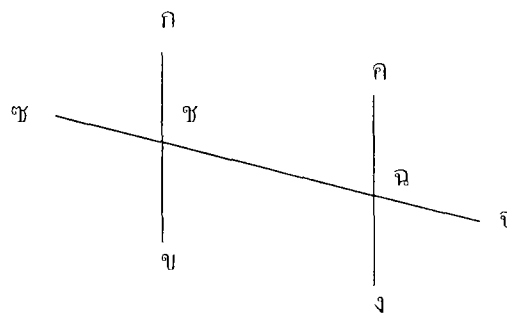
32. สายยางยาว 2.5 เมตร ราคา 125 บาท สมลักษณะซื้อสายยางยาว 15 เมตร คนขายลดให้ 30 บาท สมลักษณะต้องจ่ายเงินค่าสายยางกี่บาท

- ก. 690 บาท
- ข. 700 บาท
- ค. 720 บาท
- ง. 750 บาท

33. อ้วน และอ้อยมีอายุมากกว่าอัน 3 และ 5 ปี ตามลำดับ เมื่อรวมอายุทั้งสามคนจะได้ 41 ปี อ้อยมีอายุกี่ปี

- ก. 11 ปี
- ข. 14 ปี
- ค. 16 ปี
- ง. 21 ปี

34. จากรูปที่กำหนดให้ เงื่อนไขในข้อใดสรุปได้ว่า $กข \parallel คง$

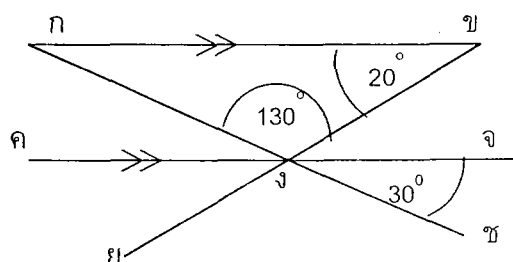


- ก. $\angle ข + \angle ก = 180^\circ$
- ข. $\angle ข + \angle ค = 180^\circ$
- ค. $\angle ก + \angle ค = 180^\circ$
- ง. $\angle ข + \angle ง = 180^\circ$

35. สวนชนิดหนึ่งกว้าง 12 เมตร ยาว 16 เมตร ต้องการทำรั้วล้อมรอบสวน โดยใช้เสาปักให้มีระยะห่างเท่ากันตลอด และใช้เสาน้อยต้นที่สุด จงหาว่าต้องใช้เสาทั้งหมดกี่ต้น

- ก. 12 ต้น
- ข. 13 ต้น
- ค. 14 ต้น
- ง. 15 ต้น

36. จากรูป กำหนดให้ $\overline{กข} \parallel \overline{คจ}$ มุมที่มีขนาดเท่ากับ กังค คือมุมในข้อใด



- ก. ขงจ
- ข. คงช
- ค. จงช
- ง. คงย

37. ตะกั่วแท่งหนึ่งหนัก $3\frac{3}{5}$ กิโลกรัม อีกแท่งหนึ่งหนัก $2\frac{2}{5}$ กิโลกรัม ส่วนตะกั่วแท่งที่สาม

หนักเป็นครึ่งหนึ่งของน้ำหนักตะกั่วแท่งแรกรวมกับตะกั่วแท่งที่สองและเพิ่มอีก $1\frac{2}{3}$

กิโลกรัม ตะกั่วแท่งที่สามหนักกี่กิโลกรัม

ก. $2\frac{4}{5}$ กิโลกรัม

ข. $5\frac{1}{3}$ กิโลกรัม

ค. $4\frac{2}{3}$ กิโลกรัม

ง. $7\frac{2}{3}$ กิโลกรัม

38. การเปรียบเทียบเศษส่วนข้อใดถูกต้อง

ก. $\frac{2}{5} < \frac{2}{9}$

ข. $\frac{2}{5} = \frac{14}{45}$

ค. $\frac{21}{22} > \frac{3}{4}$

ง. $\frac{3}{5} > \frac{3}{2}$

39. กำหนดให้ $m = b \times c$ และ $n = b \times d$ ถ้า m และ n เป็นจำนวนนับใด ๆ
 b, c และ d เป็นจำนวนเฉพาะ ห.ร.ม. ของ m และ n ได้เท่ากับข้อใด

ก. b

ข. c

ค. d

ง. cd

40. $0.25 + \square = 0.75$

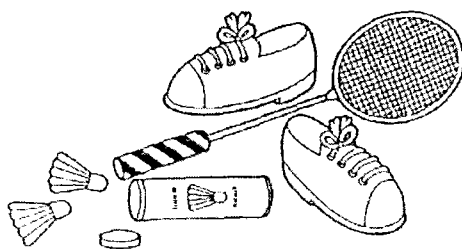
ก. $\frac{1}{2}$

ข. $\frac{1}{4}$

ค. $\frac{1}{5}$

ง. $\frac{1}{25}$

41. จากรูป จำนวนลูกแบดมินตันในกล่องมีมากกว่า 5 เท่าของจำนวนลูกแบดมินตันนอกกล่องอยู่ 1 ลูก จำนวนลูกแบดมินตันในกล่องมีกี่ลูก



- ก. 5 ลูก
ข. 8 ลูก
ค. 9 ลูก
ง. 11 ลูก

42.
$$\begin{array}{r} 2 \\ 5 \\ 5 \end{array} \begin{array}{r} 5 \\ 3 \\ + \\ 5 \end{array} = \square$$

- ก. $\begin{array}{r} 11 \\ \hline 20 \end{array}$
ข. $\begin{array}{r} 12 \\ \hline 43 \end{array}$
ค. $\begin{array}{r} 43 \\ \hline 75 \end{array}$
ง. $\begin{array}{r} 7 \\ 3 \hline 12 \end{array}$

43. $\frac{1}{3}$ ของ $\frac{3}{4}$ ของจำนวนที่ดิน 1 ไร่ 2 งาน คิดเป็นที่ดินกี่ตารางวา
(1 ไร่ = 400 ตารางวา, 1 ไร่ = 4 งาน)

- ก. 100 ตารางวา
ข. 150 ตารางวา
ค. 200 ตารางวา
ง. 250 ตารางวา

44. สายไฟเส้นหนึ่งยาว 12 เมตร ตัดออกเป็น 2 ส่วน ส่วนแรกยาว $5\frac{1}{4}$ เมตร
อีกส่วนหนึ่งยาวกี่เมตร เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใด

ก. $12 - 5\frac{1}{4} = \square$

ข. $12 \times 5\frac{1}{4} = \square$

ค. $12 \div 5\frac{1}{4} = \square$

ง. $(12 \div 2) - 5\frac{1}{4} = \square$

45. รถโดยสารคันหนึ่งแล่นจากกาญจนบุรีไปกรุงเทพฯ เมื่อถึงนครปฐมมีผู้โดยสารลง
รถ $\frac{1}{3}$ ของผู้โดยสารทั้งหมด ปรากฏว่ายังเหลือผู้โดยสารอีก 60 คน อยากทราบว่า
ผู้โดยสารลงรถที่นครปฐมกี่คน

ก. 15 คน

ข. 20 คน

ค. 30 คน

ง. 48 คน

46. แม่ค้าซื้อส้มมา $30\frac{4}{5}$ กิโลกรัม ราคา กิโลกรัมละ 6 บาท แล้วนำไปขายราคา กิโลกรัมละ

$7\frac{1}{2}$ บาท เมื่อขายหมดแม่ค้าจะได้กำไรกี่บาท เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใด

ก. $(7\frac{1}{2} - 6) \times 30\frac{4}{5} = \square$

ข. $30\frac{4}{5} \times (7\frac{1}{2} + 6) = \square$

ค. $(30\frac{4}{5} \times 6) - (30\frac{4}{5} \times 7\frac{1}{2}) = \square$

ง. $(30\frac{4}{5} \times 6) + (30\frac{4}{5} \times 7\frac{1}{2}) = \square$

47. จำนวนในข้อใดมีจำนวนนับเป็นตัวประกอบทั้งหมด 4 ตัว

ก. 30

ข. 32

ค. 35

ง. 42

48. ห.ร.ม.ของ $(5^3 \times 7)$ กับ $(3 \times 5^2 \times 7^3)$ คือจำนวนในข้อใด

ก. 5^3

ข. 3×5^2

ค. $5^3 \times 7^3$

ง. 7×5^2

49. $\left(5\frac{1}{3} \div 2\right) + 7 = \square$

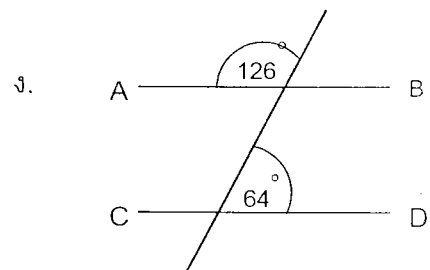
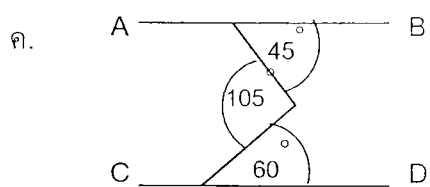
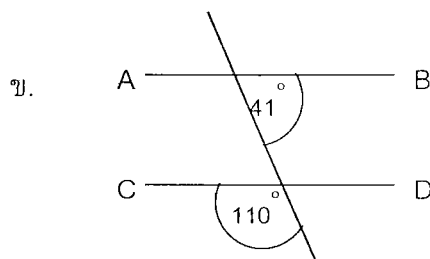
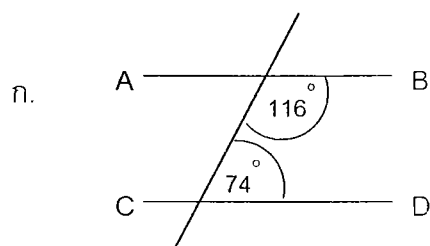
ก. 5

ข. 15

ค. $8\frac{2}{3}$

ง. $9\frac{2}{3}$

50. จากรูป จงพิจารณาว่าข้อใดที่ทำให้ $\overline{AB} \parallel \overline{CD}$



ภาคผนวก จ

หนังสือเชิญเป็นผู้เชี่ยวชาญ



บันทึกข้อความ

ส่วนราชการ บัณฑิตวิทยาลัย มศว โทร. ๕๗๒๖, ๕๖๔๖

ที่ ทม ๑๐๑๒/๔๔๘๓

วันที่ ๖ ธันวาคม ๒๕๕๓

เรื่อง ขอเชิญเป็นผู้เชี่ยวชาญ

เรียน คณบดีคณะศึกษาศาสตร์

เนื่องด้วย นางพิมพ์ ภูมิรา นิสิตระดับปริญญาโท วิชาเอกการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับอนุมัติให้ดำเนินการทำปริญญานิพนธ์เรื่อง “การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎชิลิ่งที่แตกต่างกัน ๓ แบบ” โดยมีรองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ และรองศาสตราจารย์ ภา สรีไพโรจน์ เป็นคณะกรรมการควบคุมการทำปริญญานิพนธ์ ในการนี้ บัณฑิตวิทยาลัยขอเรียนเชิญ อาจารย์ ดร.สุพร เข้มเฮง และอาจารย์ชวลิต รวยอาจิน เป็นผู้เชี่ยวชาญตรวจสอบแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่าน ได้โปรดพิจารณาให้ข้าราชการในสังกัดเป็นผู้เชี่ยวชาญให้ นางพิมพ์ ภูมิรา ซึ่งจะเป็นประโยชน์ในการพัฒนาคุณภาพการศึกษา และขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

คณบดีบัณฑิตวิทยาลัย

ที่ ทม ๑๐๑๒, ๔๒๗๔



162

บัณฑิตวิทยาลัย

มหาวิทยาลัยศรีนครินทรวิโรฒ

สุขุมวิท ๒๓ กรุงเทพฯ ๑๐๑๔๐

๒ ธันวาคม ๒๕๔๓

เรื่อง ขอเชิญเป็นผู้เชี่ยวชาญ

เรียน ผู้อำนวยการกองพัฒนา กรมการศึกษานอกโรงเรียน

สิ่งที่ส่งมาด้วย แบบทดสอบ

เนื่องด้วย นางพิมพ์ ภูมิรา นิสิตระดับปริญญาโท วิชาเอกการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับอนุมัติให้ดำเนินการทำปริญญานิพนธ์ เรื่อง “การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎซิลลิงที่แตกต่างกัน ๓ แบบ” โดยมี รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ และ รองศาสตราจารย์นิภา ศรีไพโรจน์ เป็นคณะกรรมการควบคุมการทำปริญญานิพนธ์ ในการนี้ บัณฑิตวิทยาลัยขอเรียนเชิญ ท่านเป็นผู้เชี่ยวชาญตรวจแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่าน ได้โปรดพิจารณาเป็นผู้เชี่ยวชาญให้นางพิมพ์ ภูมิรา ซึ่งจะเป็นประโยชน์ในการพัฒนาคุณภาพการศึกษา และขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

คณบดีบัณฑิตวิทยาลัย

สำนักงานคณบดีบัณฑิตวิทยาลัย

โทร. ๖๖๔-๑๐๐๐ ต่อ ๕๗๒๖, ๕๖๔๔

โทรสาร. ๒๕๘-๔๑๑๕



ที่ ทม ๑๐๑๒/๔๒๔๔

บัณฑิตวิทยาลัย

มหาวิทยาลัยศรีนครินทรวิโรฒ

สุขุมวิท ๒๓ กรุงเทพฯ ๑๐๑๑๐

๖ ธันวาคม ๒๕๕๓

เรื่อง ขอเชิญเป็นผู้เชี่ยวชาญ

เรียน ผู้อำนวยการศูนย์เทคโนโลยีการศึกษา กรมการศึกษานอกโรงเรียน

สิ่งที่ส่งมาด้วย แบบทดสอบ

เนื่องด้วย นางพิมพ์ ภูมิรา นิสิตระดับปริญญาโท วิชาเอกการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับอนุมัติให้ดำเนินการทำปริญญานิพนธ์ เรื่อง “การศึกษาลำความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎชิลิ่งที่แตกต่างกัน ๓ แบบ” โดยมี รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนันตพงษ์ และ รองศาสตราจารย์นิภา ศรีไพโรจน์ เป็นคณะกรรมการควบคุมการทำปริญญานิพนธ์ ในการนี้ บัณฑิตวิทยาลัยขอเรียนเชิญ นายขรรค์ชัย คงสนั่น นักวิชาการศึกษา ๘ เป็นผู้เชี่ยวชาญตรวจแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่าน ได้โปรดพิจารณาให้ข้าราชการในสังกัดเป็นผู้เชี่ยวชาญให้ นางพิมพ์ ภูมิรา ซึ่งจะเป็นประโยชน์ในการพัฒนาคุณภาพการศึกษา และขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

คณบดีบัณฑิตวิทยาลัย

สำนักงานคณบดีบัณฑิตวิทยาลัย

โทร. ๖๖๔-๑๐๐๐ ต่อ ๕๗๒๖, ๕๖๔๔

โทรสาร. ๒๕๘-๔๑๑๕

ที่ ทม ๑๐๑๒/ ๔๔๘๖



บัณฑิตวิทยาลัย

มหาวิทยาลัยศรีนครินทรวิโรฒ

สุขุมวิท ๒๓ กรุงเทพฯ ๑๐๑๑๐

๑ ธันวาคม ๒๕๔๓

เรื่อง ขอเชิญเป็นผู้เชี่ยวชาญ

เรียน อาจารย์ใหญ่โรงเรียนบ้านทุ่งประทุน

สิ่งที่ส่งมาด้วย แบบทดสอบ

เนื่องด้วย นางพิมพ์ ภูมิรา นิสิตระดับปริญญาโท วิชาเอกการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับอนุมัติให้ดำเนินการทำปริญญานิพนธ์เรื่อง “การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎชิลล์ที่แตกต่างกัน ๓ แบบ” โดยมี รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์ และ รองศาสตราจารย์ ภา สรีไพโรจน์ เป็นคณะกรรมการควบคุมการทำปริญญานิพนธ์ ในการนี้ บัณฑิตวิทยาลัยขอเรียนเชิญ อาจารย์ธานี เขื่อนมัน เป็นผู้เชี่ยวชาญตรวจแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่าน ได้โปรดพิจารณาให้ข้าราชการในสังกัดเป็นผู้เชี่ยวชาญให้ นางพิมพ์ ภูมิรา ซึ่งจะเป็นประโยชน์ในการพัฒนาคุณภาพการศึกษา และขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(ศาสตราจารย์ ดร.เสริมศักดิ์ วิสาลาภรณ์)

คณบดีบัณฑิตวิทยาลัย

สำนักงานคณบดีบัณฑิตวิทยาลัย

โทร. ๖๖๔-๑๐๐๐ ต่อ ๕๗๒๖, ๕๖๔๔

โทรสาร. ๒๕๘-๔๑๑๕

ภาคผนวก จ
หนังสือขอความร่วมมือเพื่อการวิจัย

ที่ ทม ๑๐๑๒/ ๔๔๖๘



166

บัณฑิตวิทยาลัย

มหาวิทยาลัยศรีนครินทรวิโรฒ

สุขุมวิท ๒๓ กรุงเทพฯ ๑๐๑๑๐

๒๗ ธันวาคม ๒๕๔๓

เรื่อง ขอความร่วมมือเพื่อการวิจัย

เรียน ผู้อำนวยการการประถมศึกษาจังหวัดกาญจนบุรี

สิ่งที่ส่งมาด้วย แบบทดสอบ และรายชื่อโรงเรียนในสังกัด

เนื่องด้วย นางพิมพ์ ภูมิรา นิสิตระดับปริญญาโท วิชาเอกการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับอนุมัติให้ดำเนินการทำปริญญานิพนธ์ เรื่อง “การศึกษาค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎชดลิ่งที่แตกต่างกัน ๓ แบบ” โดยมี รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ และ รองศาสตราจารย์นิภา ศรีไพโรจน์ เป็นคณะกรรมการควบคุมการทำปริญญานิพนธ์ ในการนี้ นิสิตมีความจำเป็นต้องเก็บข้อมูลเพื่อการวิจัย โดยขออนุญาตใช้สถานที่และขอให้นักเรียนระดับประถมศึกษาชั้นปีที่ ๖ โรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดกาญจนบุรี ทำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๖ โดยกฎชดลิ่งที่แตกต่างกัน ๓ แบบ ในระหว่างเดือนมกราคม ๒๕๔๔

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่าน ได้โปรดพิจารณาให้ นางพิมพ์ ภูมิรา ได้เก็บข้อมูลในการทำปริญญานิพนธ์ และจะกรุณามีหนังสือแจ้งไปยังโรงเรียนในสังกัดจะเป็นพระคุณอย่างยิ่ง

ขอแสดงความนับถือ

(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

คณบดีบัณฑิตวิทยาลัย

สำนักงานคณบดีบัณฑิตวิทยาลัย

โทร. ๖๖๔-๑๐๐๐ ต่อ ๕๓๒๖, ๕๖๔๔

โทรสาร. ๒๕๘-๔๑๑๕

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อนางพิมพ์ร ภูมิรา

เกิดวันที่ 28 เดือนกุมภาพันธ์

พุทธศักราช 2496

สถานที่เกิด

อำเภอศรีสำโรง จังหวัดสุโขทัย

สถานที่อยู่ปัจจุบัน

บ้านเลขที่ 7 ซอยสุขาภิบาลท่ามะกา 24

ถนนแสงชูโต ตำบลท่าเรือ อำเภอท่ามะกา

จังหวัดกาญจนบุรี 71130

ตำแหน่งหน้าที่การงานในปัจจุบัน

อาจารย์ 3 ระดับ 8

สถานที่ทำงานปัจจุบัน

โรงเรียนวัดท่าเรือ “อัสสหาะวิทยาการ”

ประวัติการศึกษา

พ.ศ. 2513

มัธยมศึกษาปีที่ 3 จากโรงเรียนสตรีศรีอยุธยา

พ.ศ. 2515

ประกาศนียบัตรวิชาการศึกษาชั้นต้น (ป.กศ.ต้น) จาก
วิทยาลัยครูสวนสุนันทา

พ.ศ. 2519

กศ.บ. (ชีววิทยา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร

พ.ศ. 2544

กศ.ม. (การวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทร
วิโรฒ ประสานมิตร