

371.26013

พ.ศ. 2539

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ
โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู

ปริญญานิพนธ์

ของ

พิมพ์อุษา เจริญยศ

22 ต.ค. 2539

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

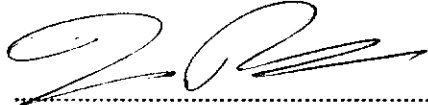
กุมภาพันธ์ 2539

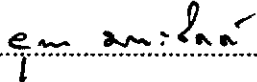
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

B. 5005A

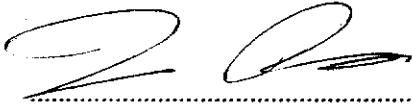
คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอก
การวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

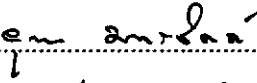
คณะกรรมการควบคุม


.....ประธาน
(รองศาสตราจารย์ ดร.นุชเชิด ภิญโญนนตพงษ์)


.....กรรมการ
(อาจารย์ยุพา มานะจิตต์)

คณะกรรมการสอบ


.....ประธาน
(รองศาสตราจารย์ ดร.นุชเชิด ภิญโญนนตพงษ์)


.....กรรมการ
(อาจารย์ยุพา มานะจิตต์)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(ผู้ช่วยศาสตราจารย์ ดร.พิตร ทองชั้น)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัย
ศรีนครินทรวิโรฒ


..... คณบดีบัณฑิตวิทยาลัย
(ดร.ศิริยุภา พูลสุวรรณ)

วันที่ 10 เดือน กุมภาพันธ์ พ.ศ. 2539

ปริญญานิพนธ์ฉบับนี้ได้รับทุนสนับสนุนการวิจัยจาก
“ศาสตราจารย์ ดร.ชวาล แพรัตกุล”
ผู้วิจัยขอโน้มรำลึกถึงพระคุณในครั้งนี้

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความสามารถและความกรุณาอย่างสูงยิ่งของรองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ และอาจารย์ยุพา มานะจิตต์ ที่ได้เอาใจใส่ให้คำปรึกษาแนะนำ ชี้แนะแนวทางตรวจแก้ไขข้อบกพร่องต่าง ๆ อย่างดียิ่ง ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้ ขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.พิตร ทองชั้น ที่ได้กรุณาให้ข้อเสนอแนะเพิ่มเติม เพื่อให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณผู้เชี่ยวชาญทุกท่านที่ช่วยประเมินความสอดคล้องระหว่างข้อคำถามแต่ละข้อ กับจุดประสงค์เชิงพฤติกรรม และอาจารย์ผู้สอนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ทุกท่านที่ช่วยในการพิจารณาคัดเลือกข้อสอบให้มีจำนวนข้อลดลง

ขอกราบขอบพระคุณสำนักงานคณะกรรมการการประถมศึกษาจังหวัดสุพรรณบุรี และสำนักงานการประถมศึกษาอำเภอเมืองสุพรรณบุรี อำเภอบางปลาม้า และอำเภอเดิมบางนางบวช ที่กรุณาอำนวยความสะดวกในการติดต่อประสานงานกับโรงเรียนที่เป็นกลุ่มตัวอย่างเป็นอย่างดี

ขอขอบพระคุณผู้อำนวยการ อาจารย์ใหญ่ และคณะครู-อาจารย์ ของโรงเรียนในกลุ่มตัวอย่างรวมทั้งนักเรียนทุกคนที่ได้ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบพระคุณอาจารย์ยุฤทธิ์ ยืนยงค์และอาจารย์ดุเชฎิ ยืนยงค์ ที่ให้ความช่วยเหลือในการวิเคราะห์ข้อมูลด้วยเครื่องคอมพิวเตอร์ และขอขอบคุณอาจารย์สาธิตา เมธนาวิณ ที่ให้ความช่วยเหลือในการวิเคราะห์ข้อมูลด้วยเครื่องคอมพิวเตอร์ และการจัดพิมพ์ตัวอักษรที่สวยงามจนสำเร็จเป็นรูปเล่ม และขอขอบคุณพี่ ๆ เพื่อน ๆ น้อง ๆ ชาววัดผลทุกคนที่ให้ความห่วงใย เป็นกำลังใจ และให้ความช่วยเหลือในการทำปริญญานิพนธ์ฉบับนี้

สุดท้ายนี้ ผู้วิจัยขอโน้มรำลึกถึงพระคุณของมารดา ครูและอาจารย์ ผู้มีพระคุณยิ่ง ที่ได้อบรมสั่งสอนให้ความเมตตา ให้ความคิดและให้กำลังใจแก่ผู้วิจัยตลอดมา

พิมพ์อุษา เจริญยศ

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	2
ความสำคัญของการศึกษาค้นคว้า	3
ขอบเขตของการศึกษาค้นคว้า	3
นิยามศัพท์เฉพาะ	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	5
ลักษณะและรูปแบบต่าง ๆ ของแบบทดสอบเลือกตอบ	5
หลักการเขียนข้อสอบแบบเลือกตอบ	8
ข้อดีของแบบทดสอบเลือกตอบ	9
การแก้ปัญหาโจทย์คณิตศาสตร์	11
ความเชื่อมั่นของแบบทดสอบ	17
การคัดเลือกข้อสอบ	21
ความยาวของแบบทดสอบ	24
งานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า	25
สมมติฐานในการศึกษาค้นคว้า	28
3 วิธีดำเนินการศึกษาค้นคว้า	29
ประชากร	29
กลุ่มตัวอย่าง	29
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	29
วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการวิจัย	30
วิธีดำเนินการเก็บรวบรวมข้อมูล	34
ลำดับขั้นการวิเคราะห์ข้อมูล	34
สถิติที่ใช้ในการวิเคราะห์ข้อมูล	36

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล	38
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	38
การเสนอผลการวิเคราะห์ข้อมูล	38
ผลการวิเคราะห์ข้อมูล	39
5 สรุป อภิปรายผล และข้อเสนอแนะ	48
ความมุ่งหมายของการศึกษาค้นคว้า	48
กลุ่มตัวอย่าง	48
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	48
วิธีดำเนินการเก็บรวบรวมข้อมูล	48
การเสนอผลการวิเคราะห์ข้อมูล	49
สรุปผลการวิจัย	49
อภิปรายผล	51
ข้อเสนอแนะ	54
บรรณานุกรม	55
ภาคผนวก	61
ประวัติย่อของผู้วิจัย	88

บัญชีตาราง

ตาราง

หน้า

1	แสดงความสัมพันธ์ระหว่างอำนาจจำแนกของข้อสอบกับค่าความเชื่อมั่นของ แบบทดสอบ ที่ประกอบด้วยข้อสอบ 100 ข้อ	22
2	แสดงรายชื่อโรงเรียนที่เป็นกลุ่มตัวอย่างที่ใช้ในการสร้างเครื่องมือและใช้ในการวิจัย	30
3	ค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ	39
4	แสดงค่าความเชื่อมั่นของแบบทดสอบ ฉบับที่ 1 ที่มีจำนวน 40 ข้อ	41
5	แสดงค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา..	42
6	แสดงค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ดุลพินิจของครู พิจารณา.....	43
7	การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบทั้ง 11 ฉบับ	44
8	การเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงทั้ง 2 วิธี เป็นรายคู่.....	45
9	การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน.....	46
10	แสดงค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม (IOC).....	63
11	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเป็นรายข้อที่ คัดเลือกไว้จำนวน 40 ข้อ (ฉบับที่ 1)	64
12	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 2)	65
13	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 3)	66
14	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 4)	67
15	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 5)	68
16	แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 6)	69

17	แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 7)	71
18	แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 8)	72
19	แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 9)	73
20	แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 10)	74
21	แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 11)	75

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1 แสดงลำดับขั้นตอนในการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า	31
---	----

บทที่ 1

บทนำ

ภูมิหลัง

การวัดผลการศึกษา เป็นขั้นตอนหนึ่งของกระบวนการเรียนรู้ที่สำคัญยิ่ง ทั้งนี้เนื่องจากการวัดผลทางการศึกษาจะเป็นตัวชี้ให้เห็นว่า กระบวนการให้การศึกษา นั้นสัมฤทธิ์ผลมากน้อยเพียงใด ในการวัดผลการศึกษานั้นต้องใช้เครื่องมือและเทคนิคต่าง ๆ ได้แก่ การทดสอบการจัดอันดับคุณภาพ การสังเกต การสัมภาษณ์ เป็นต้น สำหรับการวัดผลสัมฤทธิ์ทางการเรียน จะใช้การทดสอบโดยมีแบบทดสอบเป็นเครื่องมือในการเก็บรวบรวมข้อมูลซึ่งส่วนมากนิยมใช้แบบทดสอบปรนัยชนิดเลือกตอบ เพราะแบบทดสอบปรนัยชนิดเลือกตอบสามารถเขียนคำถามได้ครอบคลุมเนื้อหา มีความเป็นปรนัยในการตรวจให้คะแนน และมีประสิทธิภาพ ดังที่ เรมเมอร์ (Remmer, 1955 : 94-95) ได้กล่าวว่า แบบทดสอบปรนัยชนิดเลือกตอบมีข้อดีที่วัดเนื้อหาได้ครอบคลุม ให้คะแนนได้ยุติธรรม ส่วนอดัมส์ (Adams, 1964: 331) กล่าวว่าแบบทดสอบปรนัยชนิดเลือกตอบวัดความจำเกี่ยวกับข้อเท็จจริงได้ดี และสามารถสร้างให้วัดสมรรถภาพสมองด้านต่างๆ ได้ ซึ่งเหมาะที่จะใช้ทดสอบเมื่อมีผู้สอบจำนวนมาก ส่วน ขวาล แพรัตกุล (2516: 64) ให้ความเห็นว่า การที่จะวัดความสามารถของผู้เรียนให้ได้ผลที่ถูกต้องแม่นยำและเชื่อถือได้นั้น ขึ้นอยู่กับคุณภาพของแบบทดสอบเป็นสำคัญ สำหรับแบบทดสอบที่ดีนั้น ขวาล แพรัตกุล (2516 : 123-138) ได้กล่าวว่า ควรมีคุณสมบัติ 10 ประการ คือ ความเที่ยงตรง ความเชื่อมั่น ระดับความยาก พอเหมาะ มีอำนาจจำแนก เป็นปรนัย ยุติธรรม งามลึก มีประสิทธิภาพ ยั่วเย้า และจำเพาะเจาะจง ส่วนคุณสมบัติที่สำคัญมากของแบบทดสอบนั้น เสริมศักดิ์ วิชาลาภรณ์ และเอนกกุล กริแสง (2519:179) กล่าวว่า มี 3 ประการคือ ความเชื่อมั่น ความเที่ยงตรง และความเหมาะสมที่จะใช้สำหรับสมบุรณ์ ชิตพงษ์ และสำเร็จ บุญเรืองรัตน์ (2518 : 19) อนันต์ ศรีโสภา (2525 : 59) และ บุญเจิด ภิญโญอนันตพงษ์ (2525 : 128) กล่าวว่า คุณสมบัติที่มีความสำคัญที่สุดของเครื่องมือในการวัดผลการศึกษา คือ ความเชื่อมั่นและ ความเที่ยงตรง

คุณภาพของแบบทดสอบในด้านความเชื่อมั่นนั้นนับว่าสำคัญมาก ถ้าแบบทดสอบมีความเชื่อมั่นสูง แสดงว่าคะแนนที่ได้จากการสอบ มีความเชื่อถือได้มากใกล้จะเป็นคะแนนจริงของผู้สอบ แต่ถ้าแบบทดสอบมีความเชื่อมั่นต่ำ ผลการสอบก็เชื่อถือไม่ได้เพราะมีความคลาดเคลื่อนของคะแนนอยู่มาก ส่วนความเชื่อมั่นของแบบทดสอบจะสูงหรือต่ำนั้น มีองค์ประกอบที่เกี่ยวข้องอยู่หลายประการ เช่น ความยาวของแบบทดสอบ ความยากของแบบทดสอบ ความเป็น

ปรนัยของตัวข้อสอบ ลักษณะของกลุ่มที่สอบ การดำเนินการเกี่ยวกับการสอบ เป็นต้น

ความยาวของแบบทดสอบ นับว่าเป็นองค์ประกอบสำคัญที่มีอิทธิพลต่อความเชื่อมั่นอีกประการหนึ่ง แบบทดสอบที่มีจำนวนข้อมากจะมีความเชื่อมั่นสูงกว่าแบบทดสอบที่มีจำนวนข้อน้อย ซึ่งอดัมส์ (Adams.1964 : 97) ได้กล่าวถึงการเพิ่มจำนวนข้อสอบเพื่อให้ค่าความเชื่อมั่นสูงขึ้นว่าจำนวนข้อที่เพิ่มขึ้นจะต้องมีคุณภาพเหมือนข้อสอบเดิม และการเพิ่มนั้นจะต้องให้พอเหมาะสมควรกับเวลาด้วย ในกรณีที่แบบทดสอบมีความยาวมากเกินไป อาจจะไม่ทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงตามทฤษฎี เนื่องจากนักเรียนจะเหนื่อยล้า และเกิดความเบื่อหน่ายที่จะคิดตอบคำถามต่าง ๆ ทำให้นักเรียนไม่ตั้งใจทำแบบทดสอบนั้น ๆ หรืออาจเดาคำตอบเพื่อให้การทดสอบเสร็จสิ้นอย่างรวดเร็วส่งผลให้การวิจัยผิดพลาดอย่างน่าเสียดาย

ปัญหาที่กล่าวมานี้น่าจะมีการศึกษาว่าถ้าลดความยาวของแบบทดสอบแต่ละฉบับลงบ้าง และให้มีผลกระทบต่อความเชื่อมั่นน้อยสุด หรือมีผลกระทบต่อผลการสอบตามจุดมุ่งหมายของการสอบนั้นน้อยที่สุด จากการทดลองของวิเชียร เกตุสิงห์ (2523 : 240-241) ใช้วิธี “การวิเคราะห์องค์ประกอบ” โดยใช้คะแนนรายข้อมาวิเคราะห์ด้วยเหตุผลที่ว่า ในแบบทดสอบแต่ละฉบับนั้น มีข้อสอบหลาย ๆ ข้อ หรือเป็นกลุ่มที่วัดพฤติกรรม หรือความสามารถอะไรบางอย่างร่วมกันอยู่ นั่นคือ มีข้อสอบหลายข้อวัดซ้ำ ๆ กัน น่าจะตัดข้อสอบที่วัดในสิ่งเดียวกันออกไปเสียบ้าง จากผลการทดลองปรากฏว่าสามารถลดข้อสอบลงเหลือ 65 % ของจำนวนข้อสอบที่มีคุณภาพที่มีอยู่ทั้งหมด ค่าความเชื่อมั่นของแบบทดสอบลดลงเพียงเล็กน้อย แต่ประโยชน์ที่เห็นได้ชัดคือประหยัดทั้งเวลา แรงงานในการสร้างแบบทดสอบ และเวลาในการทำแบบทดสอบ

ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีความสนใจที่จะศึกษาว่าความยาวของแบบทดสอบ เมื่อลดความยาวให้สั้นลง ให้มีจำนวนน้อยข้อนั้นควรจะมีจำนวนข้อเท่าใด และเกณฑ์การคัดเลือกข้อสอบนั้น อาจจะใช้ค่า P และ r แต่ในบางครั้งค่า P และ r ก็ไปตัดข้อที่มีความสำคัญที่ครูอยากจะทดสอบออกไป ดังนั้นจุดมุ่งหมายครั้งนี้จึงต้องการศึกษาว่าจำนวนข้อที่น้อยที่สุดควรเป็นเท่าใดจึงจะทำให้ ค่าความเชื่อมั่นเท่าเดิม และลดจำนวนข้อเป็นเท่าใดจึงทำให้ค่าความเชื่อมั่นสูงสุด และการพิจารณาตัดจำนวนข้อออกโดยใช้ค่าสถิติ กับใช้ดุลพินิจของครู จะมีผลต่อค่าความเชื่อมั่นต่างกันหรือไม่

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ ระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ
2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบระหว่าง ฉบับที่มีจำนวน 40 ข้อ กับ

ฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครู

3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้จะเป็นแนวทางในการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนให้มีจำนวนข้อที่เหมาะสม โดยมีค่าความเชื่อมั่นสูงสุด

ขอบเขตของการศึกษาค้นคว้า

1. ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษา จังหวัดสุพรรณบุรี

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดสุพรรณบุรี จำนวนนักเรียน 608 คน

3. ตัวแปรที่ศึกษาค้นคว้า

3.1 ตัวแปรอิสระในการศึกษาค้นคว้าครั้งนี้ ได้แก่

3.1.1 ความยาวของแบบทดสอบ จำแนกเป็นขนาดต่าง ๆ ดังนี้

- ความยาว 40 ข้อ
- ความยาว 35 ข้อ
- ความยาว 30 ข้อ
- ความยาว 25 ข้อ
- ความยาว 20 ข้อ
- ความยาว 15 ข้อ

3.1.2 วิธีการกำหนดความยาวมี 2 เกณฑ์ คือ

- เกณฑ์ค่าสถิติ
- เกณฑ์ดุลพินิจของครู

3.2 ตัวแปรตาม คือ ค่าความเชื่อมั่นของแบบทดสอบ

4. เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้า เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 4 ตัวเลือก

จำนวน 40 ข้อที่วัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ด้านโจทย์ปัญหา ชั้นประถมศึกษาปีที่ 5 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) กระทรวงศึกษาธิการ ของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ซึ่งผู้วิจัยสร้างขึ้นโดยครอบคลุมเนื้อหา เรื่อง โจทย์ปัญหาเศษส่วน โจทย์ปัญหาคณนิยม และโจทย์ปัญหาร้อยละ

นิยามศัพท์เฉพาะ

1. ความยาวของแบบทดสอบ หมายถึง จำนวนข้อสอบที่ประกอบกันเป็นแบบทดสอบ ในการศึกษาครั้งนี้กำหนดความยาวของแบบทดสอบเป็น 6 ขนาด คือ ความยาวขนาด 40, 35, 30, 25, 20 และ 15 ข้อ ตามลำดับ
2. การลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติ หมายถึง การพิจารณาคัดเลือกข้อสอบบางข้อ ออกไปโดยใช้การพิจารณาจากค่าสถิติของข้อสอบคือ ความยากง่าย และค่าอำนาจจำแนก ที่มีความเหมาะสมตามเกณฑ์น้อยที่สุดออกไปจุดประสงค์ละ 1 ข้อ
3. การลดจำนวนข้อโดยใช้เกณฑ์ดุลพินิจของครู หมายถึง การพิจารณาคัดเลือกข้อสอบ บางข้อออกไปโดยใช้การพิจารณาจากค่าดัชนีความสอดคล้อง (IOC) ซึ่งได้จากการประเมินความ ยากของเนื้อหาข้อสอบจากผู้เชี่ยวชาญที่เป็นครูผู้สอน
4. ความเชื่อมั่นของแบบทดสอบ หมายถึง สัดส่วนของความแปรปรวนจริง ซึ่งหาได้จาก สูตรของเฟลด์ต์ (Feldt)
5. ผู้เชี่ยวชาญ หมายถึง ผู้ที่มีวุฒิปริญญาโททางการวัดผลการศึกษาและสอนวิชา คณิตศาสตร์ไม่น้อยกว่า 5 ปี จำนวน 5 คน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาครั้งนี้ ผู้วิจัยขอเสนอเนื้อหาตามลำดับ ดังนี้

1. ลักษณะและรูปแบบต่าง ๆ ของแบบทดสอบเลือกตอบ
2. หลักในการเขียนข้อสอบแบบเลือกตอบ
3. ข้อดีของแบบทดสอบเลือกตอบ
4. การแก้ปัญหาจิตวิทยาคนทดสอบ
5. ความเชื่อมั่นของแบบทดสอบ
6. การคัดเลือกข้อสอบ
 - 6.1 ความยากของข้อสอบ
 - 6.2 ค่าอำนาจจำแนก
 - 6.3 การใช้เกณฑ์ดุลพินิจของครู
7. ความยาวของแบบทดสอบ
8. งานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า

ลักษณะและรูปแบบต่าง ๆ ของแบบทดสอบเลือกตอบ

แบบทดสอบเลือกตอบเป็นแบบทดสอบที่นิยมใช้กันโดยทั่วไปในทุกวงการ เพราะแบบทดสอบเลือกตอบสามารถวัดพฤติกรรมต่าง ๆ ได้หลายด้านและหลายระดับ สามารถใช้ได้กับเนื้อหาวิชาเกือบทุกลักษณะ การตรวจให้คะแนนได้ผลคงที่ ประหยัดเวลาและแรงงาน นอกจากนี้ยังสามารถวิเคราะห์ได้ว่า ข้อใดดีหรือไม่ดี (ชวาล แพรัตกุล, 2516:164) ในการวัดผลสัมฤทธิ์นั้น สิ่งที่สำคัญประการหนึ่ง คือ ต้องการวัดผลการเรียนรู้ในระดับความรู้ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า (Gronlund, 1976:270) จะเห็นได้ว่าแบบทดสอบเลือกตอบสามารถวัดระดับความรู้เหล่านี้ได้ดี จึงถูกนำไปใช้ในการสอบทุกระดับชั้น ตั้งแต่ประถมศึกษาจนถึงระดับมหาวิทยาลัยและได้มีการพัฒนาวิธีเขียนข้อคำถามของแบบทดสอบเลือกตอบ ให้มีคำถามและตัวเลือกแตกต่างกันออกไป เพื่อจะสามารถเลือกใช้ให้เหมาะสมกับเนื้อหาเรื่องราว หรือจุดประสงค์ของเนื้อหาวิชา ตลอดจนใช้ในการวัดผลการเรียนรู้ในระดับชั้นต่าง ๆ ได้อย่างมีประสิทธิภาพ

จากการศึกษาเอกสารและงานวิจัยของนักวัดผลการศึกษาหลาย ๆ ท่าน พอสรุปได้ว่า

แบบทดสอบเลือกตอบจะประกอบด้วยส่วนที่สำคัญ 2 ส่วน คือ

1. ส่วนที่เป็นตัวปัญหา หรือคำถาม หรือตอนนำของเรื่องราว
2. ส่วนที่เป็นตัวคำตอบ หรือตัวเลือกที่สอดคล้องกับเรื่องราวในตอนนำ ซึ่งตัวเลือกที่เป็นคำตอบถูกเรียกว่า ตัวถูก และตัวเลือกที่ผิด เรียกว่า ตัวลวง

การจำแนกรูปแบบทดสอบเลือกตอบ ได้มีผู้จำแนกไว้ต่าง ๆ กันตามลักษณะดังนี้

บุญเชิด ภิญโญนนตพงษ์ (2525 : 60-63) ได้จำแนกรูปแบบของแบบทดสอบเลือกตอบไว้เป็น 2 แบบคือ

1. แบบคำถามเดียว
 - 1.1 แบบเลือกตอบข้อถูกเพียงข้อเดียว
 - 1.2 แบบเลือกตอบข้อถูกได้หลายตัวเลือก
2. คำถามแบบรวมกลุ่ม หรือแบบตัวเลือกคงที่

เมห์เรนส์ และเลห์แมนน์ (Mehrens and Lehman, 1984:153-154) และอนันต์ ศรีโสภา (2525 : 178-180) จำแนกข้อสอบเลือกตอบออกตามลักษณะแบบฟอร์มได้เป็น 4 ลักษณะ คือ

1. มีคำตอบถูกเพียงคำตอบเดียว
2. มีคำตอบดีที่สุดเพียงคำตอบเดียว
3. ประเภทการเปรียบเทียบ
4. ประเภทกลับตรงข้าม

ชวาล แพร์ตกุล (2530 : 58-384) ได้จำแนกเป็นรูปแบบใหญ่ ๆ ได้ 3 แบบ คือ

1. แบบคำถามเดียว ลักษณะของข้อสอบแต่ละข้อจะมีคำถามเดียวโดด ๆ
2. แบบคำถามรวมกลุ่ม ลักษณะคำถามประเภทนี้ จะรวมเนื้อหาบางตอนที่มีความเป็นเอกพันธ์ร่วมกันอย่างใดอย่างหนึ่ง ไว้เป็นตัวเลือกตอบแล้วเขียนคำถามเป็นชุด ๆ นักเรียนต้องใช้ความสามารถหลาย ๆ ด้านมาผสมผสานกันจึงจะสามารถตอบได้ถูกต้อง

3. แบบสถานการณ์ เป็นแบบที่ใช้วิธีการกำหนด ข้อความ ภาพ ตาราง ให้พิจารณา แล้วตั้งคำถามเกี่ยวกับข้อความ ภาพ หรือตารางที่กำหนดให้

นอกจากนี้ ชวาล แพร์ตกุล (2530 : 62-345) เสนอรูปแบบของข้อสอบเลือกตอบประเภทคำถามเดียวไว้อีก 14 รูปแบบ คือ

1. แบบคำตอบถูก แบ่งออกเป็นชนิดตัวเลือกหรือคำตอบถูกต้องตัวเดียว คำตอบดีที่สุด และคำตอบใกล้เคียง
2. แบบเพิ่มเติม เป็นคำถามที่ดัดแปลงจากข้อสอบเติมคำ ซึ่งแบ่งได้เป็น 2 ชนิด คือ

ชนิดให้เติมแห่งเดียว และเติม 2 แห่ง

3. แบบเปลี่ยนแทน เป็นคำถามที่ให้เปลี่ยนแปลงหรือปรับปรุงข้อความเดิมให้เป็นรูปใหม่โดยเปลี่ยนคำบางคำ บางวลี หรือบางประโยค แบ่งได้เป็น 3 ชนิด คือ ชนิดเปลี่ยนแปลงปรับปรุงและเปลี่ยนแทน

4. แบบคำตอบไม่จำกัด เป็นแบบที่ผู้สอบต้องพิจารณาตัวเลือกทุกตัวว่าตัวใดถูกหรือผิด และขีดคำตอบไปตามนั้นทุกตัวเลือก

5. แบบคำตอบรวม ดัดแปลงมาจากแบบคำตอบไม่จำกัด แทนที่จะตอบทีละตัวเลือก ก็ใช้คำตอบรวมเลย แบบคำตอบรวมนี้มี 2 ชนิด คือ ชนิดคำตอบผสม และคำตอบคู่

6. แบบคำตอบไม่สมบูรณ์ เป็นแบบที่ให้ผู้สอบตอบ โดยย่อตัวเลือกเป็นเพียงตัวอักษรหรือบอกตัวเลขเพียงบางหลักเท่านั้น หรือจัดให้ตัวเลือกอยู่ในรูปคำผสมหรืออัตราส่วน หรือเปลี่ยนรูปคำตอบให้เป็นอย่างอื่น คำถามประเภทนี้มี 2 ชนิด คือ ชนิดคำตอบย่อ และชนิดคำตอบไม่สำเร็จ

7. แบบคำถามนิเสธ วัดความสามารถของผู้สอบในบทกลับ มี 4 ชนิด คือ ชนิดตำแหน่งผิดประเภทของความผิด ตรงข้าม และคำตอบผิด

8. แบบเรียงอันดับ มี 5 ชนิด คือ ชนิดลำดับเรื่องราว ลำดับเวลา ลำดับคุณลักษณะ ลำดับวิธีการ และลำดับเหตุผล

9. แบบอนุกรม มีตัวอย่างให้ดูก่อน แล้วให้ผู้สอบค้นหากฎเกณฑ์หรือระบบของตัวอย่างนั้นว่าเป็นอย่างไร เพื่อหาคำตอบมี 2 ชนิด คือ ชนิดต่ออนุกรม และชนิดอนุกรมสัมพันธ์

10. แบบขาดเกิน เป็นคำถามที่ให้ผู้สอบวินิจฉัยความสมบูรณ์ของเรื่องราว มี 3 ชนิด คือ ชนิดขาด เกิน และเพียงพอ

11. แบบสัมพันธ์เป็นคำถามที่ให้หาความสัมพันธ์ระหว่างของสองสิ่งเป็นอย่างน้อยว่ามีความสัมพันธ์กันอย่างไร มี 3 ชนิด คือ ชนิดสาเหตุและผล อุปมาอุปมัย และเชื่อมโยง

12. แบบหลักการรวม มี 3 ชนิด คือ ชนิดรวมความหมาย สรุปความหมาย และขยายความหมาย

13. แบบตรรกวิทยา เป็นคำถามที่ดัดแปลงมาจากการหาเหตุผลตามหลักตรรกศาสตร์ มี 4 ชนิด คือ ชนิดทวินัย (2 ความหมาย) ชนิดปฏินัย (มีความหมาย) ชนิดอนุมัย (บทแทรก) และชนิดยุตินัย (หาข้อยุติ)

14. แบบรูปภาพ เป็นคำถามที่ช่วยเกิดความเข้าใจได้รวดเร็ว และสามารถลดความได้เปรียบในการอ่านได้น้อยลง

หลักการเขียนข้อสอบแบบเลือกตอบ

การเขียนข้อสอบแบบเลือกตอบให้ดี จำเป็นต้องเขียนให้ดีทั้งตัวคำถามและตัวเลือกซึ่ง ขวาล แพทย์กุล (2516 : 166-190) ได้กล่าวถึงการเขียนข้อสอบแบบเลือกตอบ พอสรุปได้ดังนี้

ส่วนที่เป็นคำถามหรือตัวนำ

1. เขียนคำถามให้เป็นประโยคคำถามสมบูรณ์ เว้นแต่ข้อคำถามนั้น มีจุดมุ่งหมายเพื่อให้ผู้ตอบเติมคำหรือต่อความหมาย
2. ถามให้ตรงตามเนื้อหาที่ต้องการ เพื่อให้ได้คำถามที่เป็นตัวแทนที่ดีสามารถวัดเนื้อหาที่ต้องการจะวัดได้อย่างแท้จริง
3. ถามให้ตรงจุดและชัดเจน เพื่อให้เด็กเข้าใจคำถามได้ตรงกัน และรู้ว่าต้องตอบในแง่ใด
4. คำถามควรกระชับรัดกุมไม่ใช้คำฟุ่มเฟือย และมีความชัดเจน
5. ใช้ภาษาให้เหมาะสมกับระดับผู้สอบ
6. ใช้คำถามที่เราให้ผู้ตอบใช้ความคิด ไม่ควรถามเฉพาะความจำตามตำรา หรือความจำที่เป็นรายละเอียดเกินความจำเป็น
7. พยายามหลีกเลี่ยงคำถามที่เป็นปฏิเสธ แต่ถ้าต้องการจะใช้ประโยคปฏิเสธ ควรขีดเส้นใต้เพื่อเน้นข้อความที่ต้องการให้เด่นชัด
8. ข้อคำถามหนึ่งควรถามเรื่องเดียว
9. ไม่ควรถามสิ่งที่เด็กท่องจำคล่องปาก
10. อย่าใช้คำถามข้อแรก ๆ แนะนำคำตอบข้อหลัง ๆ หรือข้อหลัง ๆ แนะนำคำตอบข้อแรก ๆ

ส่วนที่เป็นตัวเลือก

1. ตัวเลือกต้องมีความเป็นเอกพันธ์ คือ เป็นเรื่องราวหรือประเภทเดียวกันมีทิศทางเดียวกัน และมีโครงสร้างที่สอดคล้องกัน
2. ตัวเลือกควรอิสระจากกัน
3. ใช้ภาษาที่รัดกุมชัดเจน
4. ทั้งตัวถูกและตัวผิด ต้องถูกหรือผิดตามหลักวิชา
5. ควรมีตัวเลือก 4-5 ตัว
6. ความยาวของตัวเลือกควรเป็นระบบ
7. ใช้ตัวเลือกปลายเปิดให้เหมาะสม
8. หลีกเลี่ยงการแนะนำคำตอบซึ่งอาจเกิดขึ้นจาก ใช้คำขยายไม่เหมาะสม คำตอบมีคำ

ซ้ำกับคำถามหรือเกี่ยวข้อกัน คำตอบไม่กระจายอาจมีตัวถูกซ้ำๆที่หรือผลัดเวียนไปเป็นจังหวะ ๆ คำตอบหรือตัวถูกถูกหรือผิดเด่นชัดเกินไป คำตอบใช้ศัพท์หรือภาษาแปลกกว่าตัวอื่น

การเขียนข้อสอบที่ดีจึงต้องเขียนให้ชัดเจน กระทัดรัด อ่านแล้วเข้าใจได้ตรงกัน ตามได้ตรงเนื้อหา เหมาะสมกับระดับผู้สอบ และหลีกเลี่ยงการแนะนำคำตอบ ข้อสอบที่ไม่ดีอาจทำให้นักเรียนที่มีความรู้ตอบไม่ได้ เนื่องจากไม่เข้าใจคำถาม หรือนักเรียนที่ไม่มีความรู้อาจทำข้อสอบได้ เนื่องจากมีการแนะคำตอบไว้ และถ้าเป็นเช่นนั้นจะทำให้ผลของการสอบเชื่อถือไม่ได้

ข้อดีของแบบทดสอบเลือกตอบ

ชวาล แพร์ตกุล (2516 : 164) ได้กล่าวไว้ว่าในจำพวกข้อสอบที่ให้ตอบสั้น ๆ ด้วยกันว่า ชนิดที่ดีที่สุดก็คือแบบทดสอบเลือกตอบ เพราะคำถามแบบเลือกตอบมีหลายประการ เช่น ไม่ทำให้เกิดปัญหาสองแง่ได้ง่าย ๆ สามารถลดอัตราเดาลงได้มาก ใช้คำถามสำหรับวัดสมรรถภาพด้านต่าง ๆ ได้อย่างกว้างขวาง

เสริมศักดิ์ วิชาลาภรณ์ และเอนกกุล กริแสง (2517 : 22-123) ได้กล่าวว่าแบบทดสอบเลือกตอบมีข้อดีดังนี้

1. วัดได้รอบด้าน หมายถึง สามารถวัดพฤติกรรมการเรียนรู้ตั้งแต่กระบวนการแก้ปัญหาขั้นต้น เช่น ความรู้ความเข้าใจจนกระทั่งกระบวนการแก้ปัญหาขั้นสูง เช่น การวิเคราะห์ การสังเคราะห์ และการประเมินผล เป็นต้น

2. ข้อสอบแบบเลือกตอบมีโอกาสเดาถูกน้อย เป็นผลทำให้ข้อสอบมีความเชื่อมั่นสูงกว่าข้อสอบแบบถูก-ผิด ด้วยเหตุนี้ทำให้ข้อสอบมาตรฐานเป็นรูปแบบเป็นแบบทดสอบเลือกตอบเสียเป็นส่วนใหญ่

3. แบบทดสอบเลือกตอบสามารถวัดพฤติกรรมได้หลายด้าน เช่น วัดการระลึกได้ การนำเอากฎเกณฑ์ไปใช้ การหาข้อสรุป เป็นต้น ผลจากการทดสอบสามารถวัดพฤติกรรมได้หลายด้าน ทำให้ข้อสอบมีความเที่ยงตรง คือสามารถวัดในสิ่งที่เราต้องการจะวัด

4. แบบทดสอบเลือกตอบอาจจะออกให้ง่ายหรือยากก็ได้ สามารถใช้ทดสอบกับนักเรียนเกือบทุกระดับชั้น

5. แบบทดสอบเลือกตอบสามารถออกข้อสอบได้ทุกวิชา

6. แบบทดสอบแบบเลือกตอบเหมาะที่จะใช้ในการวิเคราะห์ข้อสอบ สามารถหาค่าความยากง่ายของข้อสอบ หาค่าอำนาจจำแนกของข้อสอบ แล้วนำผลจากการวิเคราะห์ไปปรับปรุงคุณภาพข้อสอบให้มีประสิทธิภาพดียิ่งขึ้น

7. แบบทดสอบแบบเลือกตอบนี้ มีความเที่ยงตรง มีความเป็นปรนัยในการตรวจให้

คะแนน และสามารถตรวจด้วยเครื่องคอมพิวเตอร์ได้อีกด้วย

8. แบบทดสอบแบบเลือกตอบดีกว่าแบบทดสอบเติมคำ ในแง่ที่ทำให้ปัญหาเรื่องความกำกวมหมดไปได้ เพราะมีคำตอบให้เลือก และเลือกข้อที่ดีที่สุดหรือถูกที่สุด

9. สามารถใช้แผนผัง รูปภาพ กราฟ เป็นตัวปัญหาได้ง่าย ทำให้นักเรียนไม่เบื่อเวลาทำข้อสอบ

10. ผู้ออกข้อสอบไม่ต้องกังวลกับลักษณะที่เป็นเอกพันธ์ ของเนื้อหาเหมือนข้อสอบแบบจับคู่ เพราะข้อสอบแบบเลือกตอบสามารถออกให้ข้อความสั้นสุดในตัวเองได้ในข้อสอบแต่ละข้อ

11. เนื่องจากข้อสอบแบบนี้จำเป็นต้องมีตัวหลอกส่อ จึงทำให้สามารถวินิจฉัยได้ว่านักเรียนมีความบกพร่องหรือไม่เข้าใจในวิชาที่เรียนมาอย่างไร

ฮอฟกิน และแอนทีส (Hopkins and Antes, 1979 : 110) กล่าวว่า แบบทดสอบเลือกตอบมีส่วนดีคือ

1. วัดได้ทุกเนื้อหาในระดับพฤติกรรมที่ต่างกัน
2. ลดการเดาถ้าใช้ 4 ตัวเลือก หรือ 5 ตัวเลือก
3. มีค่าความเชื่อมั่นดีกว่าแบบทดสอบถูก-ผิด
4. ความยากของแต่ละข้อสามารถควบคุมได้โดยเปลี่ยนตัวเลือก
5. ตัวเลือกในแบบทดสอบเลือกตอบ สามารถวินิจฉัยเด็กได้ว่า บกพร่องเรื่องอะไร

เมห์เรนส์และเลห์แมนน์ (Mehrens and Lehmann, 1984 : 154-155) กล่าวว่า แบบทดสอบเลือกตอบมีส่วนดี ดังนี้

1. ตรวจให้คะแนนได้รวดเร็วและถูกต้อง โดยการใช้เครื่องหรือใช้บุคคลอื่น ๆ
2. มีประสิทธิภาพมากกว่าแบบทดสอบถูก-ผิด และแบบทดสอบจับคู่
3. ความยากของแบบทดสอบ สามารถควบคุมได้ โดยการเปลี่ยนแปลงตัวเลือก
4. เทียบกับแบบทดสอบถูก-ผิด แบบทดสอบเลือกตอบเดาได้น้อยกว่า
5. ให้ข้อมูลเกี่ยวกับการวินิจฉัยความรู้ของเด็กได้
6. รายชื่อของแบบเลือกตอบมีความเชื่อมั่นสูงกว่าแบบทดสอบถูก-ผิด

ไอเคน (Aiken, 1987 : 44) ได้ศึกษางานวิจัยต่าง ๆ เกี่ยวกับแบบทดสอบเลือกตอบสรุปความสำคัญได้ ดังนี้

1. แบบทดสอบเลือกตอบมีความสามารถรอบตัวสามารถวัดได้ทั้งจุดประสงค์ธรรมดา และจุดประสงค์ที่ซับซ้อน พร้อมทั้งใช้ได้กับทุกระดับชั้น และใช้ได้กับทุกกลุ่มบุคคล
2. สามารถวัดได้ครอบคลุมความสามารถหลาย ๆ ด้านได้มากกว่า ข้อสอบแบบอัตนัย

และข้อสอบปรนัยแบบอื่น ๆ

3. เค้าได้ถูกต้องน้อยกว่าแบบทดสอบถูก-ผิด
4. มีความเป็นปรนัยในการตรวจให้คะแนน และสามารถทำได้รวดเร็ว
5. สามารถนำมาวิเคราะห์ได้ง่ายและสะดวก
6. สามารถวินิจฉัย ความรู้ความสามารถของผู้ตอบได้จากการเลือกตัวเลือกของแบบ

ทดสอบ

7. มีความเชื่อมั่นสูงกว่าแบบทดสอบอื่น ๆ เพราะมีความเป็นปรนัย และมีผลจากการ
เค้าถูกน้อย

จากการศึกษาความคิดเห็นของนักวัดผลการศึกษาหลาย ๆ ท่านจะเห็นว่า แบบทดสอบ
เลือกตอบ (Multiple-Choice Test) หมายถึงแบบทดสอบที่แต่ละข้อแบ่งออกเป็น 2 ส่วน คือ ส่วนที่
เป็นคำถาม (Item) และส่วนที่เป็นคำตอบ (Options) โดยส่วนที่เป็นคำตอบจะมีคำตอบไว้ให้เลือก
ตอบ โดยผู้สอบจะต้องพิจารณาว่าคำตอบใดเป็นข้อที่ถูกต้องที่สุดเพียงคำตอบเดียว จากตัวเลือก
4 ตัว มีตัวลวง (Distracters) ปนคละอยู่ 3 ตัว ดังนั้นผู้วิจัยจึงใช้วิธีการทดสอบวิชาคณิตศาสตร์
ฉบับโจทย์ปัญหา ด้วยแบบทดสอบเลือกตอบ ซึ่งเป็นวิธีที่เป็นปรนัยมีประสิทธิภาพ นำไปใช้
สะดวก ประหยัดทั้งแรงงานและเวลา

การแก้ปัญหาคณิตศาสตร์

บลูม (Bloom, 1971:81) ได้กล่าวถึงความมุ่งหมายของวิชาคณิตศาสตร์ว่าในด้านการ
พัฒนาสติปัญญา (Cognitive domain) นั้น เป็นการพัฒนาทางความรู้ (Knowledge) ความสามารถ
ทางสติปัญญา (Intellectual Ability) และได้แบ่งโครงสร้างของแบบทดสอบวิชาคณิตศาสตร์เป็น 5
ด้านคือ

1. ความรู้เกี่ยวกับความคิดรวบยอด (Knowledge of concept) เช่น กระบวนการทาง
คณิตศาสตร์ ความจริง และหลักการ
2. ทักษะในการคำนวณ (Computational skill)
3. ทักษะในการแก้ปัญหา (Problem solving skill)
4. ความรู้เกี่ยวกับโครงสร้างทางคณิตศาสตร์ และการพิสูจน์ที่มีเหตุผล
5. ทักษะในการสื่อความหมายในทางคณิตศาสตร์

ซาลิวสกี (Zalewski, 1987:2804-A) ได้ศึกษาองค์ประกอบในการแก้ปัญหาวทาง
คณิตศาสตร์พบว่าสิ่งที่เป็นองค์ประกอบมีดังนี้

1. ความสามารถในการเข้าใจสัญลักษณ์

2. ความสามารถในการจัดกระทำ
3. ความสามารถในการอ่านและตีความ
4. ความคิดรวบยอดทางคณิตศาสตร์ (Mathematical concepts)
5. ทักษะในการคำนวณ (Computation skill)

เกี่ยวกับขั้นตอนการแก้ปัญหาทางคณิตศาสตร์ ทบวงมหาวิทยาลัย (2524 : 147) ได้เสนอแนะขั้นตอนในการแก้ปัญหาไว้ ดังนี้

1. ทำความเข้าใจกับปัญหาโดยพิจารณาแยกแยะระหว่างสิ่งที่กำหนดให้และให้สิ่งที่ต้องการหา
2. จากสิ่งที่กำหนดให้ แปลงให้อยู่ในรูปที่เห็นชัดขึ้น โดยมีแผนภาพหรือรูปภาพประกอบ
3. จากแผนภาพที่สร้างขึ้น พิจารณานหาเงื่อนไขอื่น ๆ เพิ่มเติมอาศัยนิยามคุณสมบัติทฤษฎีต่าง ๆ ที่เคยเรียนมาแล้ว
4. ในบรรดาสิ่งที่กำหนดให้กับเงื่อนไขที่ได้เพิ่มเติมมีความเกี่ยวข้องกันอย่างไร หรือมีเงื่อนไขอะไรเพิ่มเติมอีกหรือไม่
5. คิดหาวิธีการแก้ปัญหา โดยพิจารณาถึงปัญหาที่คล้ายคลึงและทดลองเลียนแบบ วิธีการแก้ปัญหา หรืออาจวิเคราะห์จากสิ่งที่โจทย์ต้องการ พิจารณาเงื่อนไขใดอีกบ้าง จึงจะได้สิ่งที่ต้องการ และเงื่อนไขนั้นจะพบได้อย่างไร
6. เรียบเรียงจัดลำดับขั้นตอนในการแก้ปัญหา
7. ทดสอบคำตอบว่าสมเหตุสมผลหรือไม่
8. ถ้าการแก้ปัญหาทำได้หลายวิธี เลือกวิธีที่สั้นและง่ายที่สุด

จอห์นสัน (ทบวงมหาวิทยาลัย. 2524:141; อ้างอิงมาจาก Johnson. n.d.) กล่าวว่า การแก้ปัญหาต้องใช้กระบวนการทางสมองที่ซับซ้อน ต้องอาศัยความรู้พื้นฐาน มโนคติและทักษะที่เกี่ยวข้องกับปัญหานั้น ๆ การที่นักเรียนมีความสามารถในการแก้ปัญหาได้ นักเรียนควรได้รับการฝึกฝนให้มีความรู้พื้นฐาน ดังต่อไปนี้

1. มีความรู้เกี่ยวกับเนื้อหา มีความเข้าใจมโนคติและทักษะในเนื้อหาที่เกี่ยวข้องกับปัญหานั้น
2. มีความสามารถในการอ่าน การแปลความ การตีความ และการขยายความ
3. มีความสามารถในการแปลงข้อความให้เป็นสัญลักษณ์หรือแผนภาพ
4. มีความสามารถในการวิเคราะห์หาความเกี่ยวข้องระหว่างข้อมูลที่มีอยู่กับ

ประสบการณ์เก่า

5. มีความสามารถในการจัดระบบข้อมูล จัดลำดับขั้นตอน การวิเคราะห์หารูปแบบและการหาข้อสรุป

ทบทวนมหาวิทยาลัย (2524 : 99) ได้เสนอแนะลักษณะการแสดงออกของพฤติกรรมทางคณิตศาสตร์ ดังนี้

1. ทักษะการคิดคำนวณเป็นความสามารถในการคำนวณได้อย่างคล่องแคล่ว แม่นยำ หรือแทนค่าสัญลักษณ์ต่าง ๆ ได้อย่างรวดเร็วถูกต้อง ในเวลาที่กำหนดให้ เช่น การบวกลบ คูณหาร และการแก้สมการอย่างง่าย ๆ ไม่ซับซ้อน และการแทนค่าตัวแปรต่าง ๆ

2. การให้เหตุผลทางคณิตศาสตร์ เป็นความสามารถในการตีความ ขยายความ สรุปผล เปรียบเทียบบอกขั้นตอนก่อนหลัง เช่น สามารถบอกว่าคู่อันดับใดสอดคล้องกับสมการที่กำหนดให้ได้

3. การแก้ปัญหาทางคณิตศาสตร์ เป็นความสามารถในการแก้โจทย์ปัญหาต่าง ๆ ทั้งที่เคยพบเห็นและไม่เคยพบเห็นมาก่อน โดยอาศัยการคิดคำนวณ การให้เหตุผลและมโนคติต่าง ๆ ประมวลเข้าด้วยกันในการแก้ปัญหาต่าง ๆ ตลอดจนความสามารถในการติดตามการพิสูจน์

สำหรับการสร้างข้อสอบวิชาคณิตศาสตร์ ทฤษฎีของ เธอร์สโตน (Thurstone. 1947:121) ได้กล่าวว่า ความสามารถด้านตัวเลขหรือด้านคณิตศาสตร์นั้น คือ สมรรถภาพในการที่จะเข้าใจ และสามารถคำนวณตัวเลข โดยใช้พื้นฐานเบื้องต้น ผู้ที่มีความชำนาญคล่องแคล่วแม่นยำมาก ก็เป็นผู้ที่มีความสามารถในตัวเอง ลักษณะข้อสอบประเภทนี้มีโครงสร้าง ดังนี้

1. ข้อสอบประเภททักษะ เป็นการวัดความคล่องแคล่วในการคำนวณ

2. ข้อสอบคณิตศาสตร์เหตุผล เป็นการวัดสมรรถภาพ ด้านความคิดรวบยอดเชิงคณิตศาสตร์ซึ่งเกี่ยวข้องกับการย่อหรือการขยายกฎเกณฑ์ หลักการ วิธีการ และมโนภาพในความคิดแบบนามธรรม นั่นคือ เมื่อเด็กได้เรียนรู้กฎเกณฑ์ต่าง ๆ แล้วจะสามารถขยายความหมายนั้นไปในรูปภาษา หรือสัญลักษณ์ให้กว้างขวางจากเดิมได้หรือไม่

3. ข้อสอบประเภทปัญหาโจทย์ เป็นการวัดสมรรถภาพด้านการแก้ปัญหาแบบทดสอบประเภทนี้ประกอบด้วย คำถามโจทย์ปัญหาทำนองเดียวกับโจทย์คณิตศาสตร์ทั่วไป แต่ลักษณะของคำถามมุ่งที่จะวัดความสามารถในการแปรรูปให้หาความสัมพันธ์ของตัวเลข จำนวนและการใช้เทคนิคหรือวิธีการแก้ปัญหาต่างๆ นอกจากนี้เป็นความสามารถการจำสูตรและกฎเกณฑ์ต่าง ๆ แล้วต้องเอาตัวเลขจากโจทย์ไปแทนค่าในสูตร

สำหรับ ล้วน สายยศ (2518 : 26-30) ได้ให้ข้อเสนอแนะเกี่ยวกับการออกข้อสอบ

คณิตศาสตร์ว่าควรเน้น 3 ประการ คือ คณิตศาสตร์ทักษะ คณิตศาสตร์เหตุผล และคณิตศาสตร์
 โจทย์ปัญหา เพื่อจะได้รู้จุดบกพร่องในการคิดคณิตศาสตร์ของเด็กเมื่อเด็กบกพร่องด้านใดจะได้
 สร้างบทเรียนให้สอดคล้องกับปัญหาที่เกิดขึ้น เป็นการแก้ปัญหาตั้งแต่จุดเริ่มต้น

1. คณิตศาสตร์ทักษะ ในที่นี้หมายถึงความคล่องแคล่วในการบวก ลบ คูณ หาร ถอด
 ราก หรือทำเลขเศษส่วนได้อย่างรวดเร็ว ถูกต้อง ในเวลาที่กำหนดให้

2. คณิตศาสตร์เหตุผล การออกข้อสอบคณิตศาสตร์แบบนี้ จะถามเน้นในเรื่องวิธีการ
 หลักการ การแปลความ การตีความ การขยายความ การไล่เลียงหาเหตุผล การเปรียบเทียบ การ
 คิดริเริ่มสร้างสรรค์ มโนภาพ ขั้นตอนการพิสูจน์ และการประเมินค่า

3. คณิตศาสตร์โจทย์ปัญหา การเขียนข้อสอบคณิตศาสตร์ แบบโจทย์ปัญหานั้น มุ่ง
 ตรวจสอบผลสัมฤทธิ์สุดท้ายในการคิด นั่นคือ สร้างโจทย์ปัญหาทางคณิตศาสตร์ขึ้นมา แล้วหา
 ว่าผลลัพธ์เป็นเท่าไรนั่นเอง วิธีการหรือเหตุผลจะเป็นอย่างไร ข้อสอบลักษณะนี้ไม่สนใจจะสนใจ
 แต่คำตอบสุดท้ายจากการแก้ปัญหาเมื่อมีสถานการณ์ใด ๆ ขึ้นเท่านั้น

ขวาล แพร์ตกุล (2517 : 46-47) ได้ชี้ให้เห็นว่า โครงสร้างของแบบทดสอบคณิตศาสตร์
 ในปัจจุบันนิยมแยกคำถามออกเป็น 3 ฉบับคือ แบบทดสอบสำหรับวัดทักษะหรือความคล่อง
 แคล่วแม่นยำในการคำนวณตัวเลข แบบทดสอบสำหรับวัดความสามารถในการแก้โจทย์ปัญหา
 และแบบทดสอบสำหรับวัดความคิดรวบยอดหรือที่เรียกว่าแบบทดสอบวัดเหตุผลเชิงคณิตศาสตร์
 ซึ่งมีรายละเอียดดังนี้

1. แบบทดสอบประเภททักษะ แบบทดสอบชนิดนี้ต้องการวัดความคล่องแคล่วแม่นยำ
 ในการบวกลบ คูณ หาร ตัวเลข ซึ่งเป็นสมรรถภาพพื้นฐานในการเรียนคณิตศาสตร์ เทียบได้กับ
 ความสามารถในการสะกดคำของวิชาภาษาไทย ฉะนั้นโครงสร้างของแบบทดสอบชนิดนี้จึงไม่ต้อง
 การให้อธิพวลของภาษาเข้าเกี่ยวข้องด้วยเลย จัดว่าเป็นการวัดความคล่องแคล่วในการ
 คำนวณตัวเลขโดยเฉพาะ

2. แบบทดสอบประเภทโจทย์ปัญหา แบบทดสอบชนิดนี้ประกอบด้วยคำถามประเภท
 โจทย์ต่าง ๆ ทำนองเดียวกับโจทย์คณิตศาสตร์ทั่ว ๆ ไป แต่ลีลาคำถามมุ่งที่จะวัดความสามารถ
 ในการแปลความหมาย และการให้หาความสัมพันธ์ของตัวเลขจำนวน กับวิธีการแก้ปัญหาต่าง ๆ
 เป็นสำคัญ คำถามประเภทโจทย์ปัญหานี้ มิได้หมายถึงคำถามที่ให้นักเรียนนำตัวเลขไปแทนค่าใน
 สูตร โดยตรง ดังเช่น ในเรื่องการหาดอกเบี้ยหรือให้หาพื้นที่ของวงกลมเลย เพราะคำถามชนิดนี้ไม่
 มีลักษณะเป็นปัญหา แต่เป็นการวัดความสามารถในการจำสูตรและกฎเกณฑ์ต่าง ๆ แล้วก็เอาตัว
 เลขจากโจทย์ไปแทนค่าในสูตรเท่านั้นเอง

3. แบบทดสอบประเภทเหตุผล แบบทดสอบชนิดนี้ ต้องการวัดสมรรถภาพด้านความคิดรวบยอดเชิงคณิตศาสตร์ซึ่งเกี่ยวข้องกับหลักการ วิธีการ และมโนภาพ ในความคิดแบบนามธรรมนั่นคือ ต้องการวัดว่าเมื่อเด็กได้เรียนรู้กฎเกณฑ์ต่าง ๆ แล้วเขาจะสามารถขยายความหมายนั้นออกเป็นในรูปของภาษา หรือสัญลักษณ์ให้กว้างขวางจากเดิมได้หรือไม่ แบบทดสอบประเภทนี้จะช่วยให้ครูอาจารย์ได้ทราบว่าหลังจากเรียนเนื้อหาต่าง ๆ แล้ว นักเรียนเกิดความคิดในเรื่อง ความสัมพันธ์ของปริมาณเหล่านั้นหรือไม่ หรือเพียงแค่จำสูตรและวิธีการตามตำราได้เท่านั้น ไม่มีจินตนาการในกฎเกณฑ์เหล่านั้นเลย

จากที่กล่าวมาทั้งหมดนี้ จะเห็นว่า องค์ประกอบของวิชาคณิตศาสตร์และลักษณะของข้อสอบคณิตศาสตร์ มักใช้ทักษะ 3 ประการด้วยกันคือ ทักษะการคำนวณ ความคิดรวบยอดทางคณิตศาสตร์และ การแก้ปัญหาทางคณิตศาสตร์ เลอ บลองค์ (Le Blanc, 1977:16) ให้ความเห็นเพิ่มเติมว่า ในจำนวนองค์ประกอบทั้งสามนั้น การแก้ปัญหาทางคณิตศาสตร์นับได้ว่าเป็นเป้าหมายสูงสุดของการสอนคณิตศาสตร์ และเป้าหมายสูงสุดของการสอนแก้ปัญหานั้น มุ่งที่กระบวนการในการแก้ปัญหา ซึ่งตรงกับความคิดเห็นของเวล (Whirl, 1973:551) ที่เสนอในบทความเรื่อง กลวิธีการแก้ปัญหาว่า คำตอบของโจทย์ปัญหาไม่มีความสำคัญเท่ากับวิธีการที่ใช้ในการหาคำตอบเลย วิธีการที่ดีที่ใช้ในการแก้ปัญหานั้นเป็นวัตถุประสงค์อันสำคัญยิ่งในการเรียนคณิตศาสตร์ ส่วนกรุลิค (Kulik, 1977:650) กล่าวว่า การแก้ปัญหานั้นเป็นส่วนหนึ่งของการเรียนคณิตศาสตร์ ครูจะต้องสร้างสถานการณ์เพื่อให้ผู้เรียนได้เผชิญกับปัญหา เพื่อให้เขาได้ค้นหาและพินิจพิจารณาให้ได้มาซึ่งคำตอบ บางครั้งอาจจะไม่ได้คำตอบตามที่เขาวางก็ได้ แต่นั่นไม่ใช่สิ่งสำคัญเท่ากับการที่ผู้เรียนได้ฝึกทักษะ หรือการพินิจพิจารณา หมายความว่า วิธีการค้นหาคำตอบของปัญหาเป็นสิ่งสำคัญมากกว่าคำตอบในการแก้ปัญหา และเวสต์ (West, 1977:57-58) กล่าวว่านักเรียนที่สามารถบอกคำตอบของโจทย์ปัญหาทางคณิตศาสตร์ได้ถูกต้อง อาจจะไม่สามารถบอกกระบวนการคิดทางคณิตศาสตร์ที่ถูกต้องก็ได้ นั่นคือ นักเรียนที่คิดด้วยวิธีการที่ผิด อาจจะได้คำตอบที่ถูกต้องก็ได้ และนอกจากนี้เขายังชี้ให้เห็นถึงสาเหตุที่นักเรียนไม่สามารถจะทำข้อสอบที่เป็นโจทย์ปัญหาทางคณิตศาสตร์ได้ถูกต้องไว้ 3 ประการคือ

1. นักเรียนไม่สามารถจะเข้าใจในข้อความที่เป็นโจทย์ปัญหา
 2. นักเรียนไม่สามารถเปลี่ยนโจทย์ปัญหานั้น มาเป็นประโยคสัญลักษณ์ทางคณิตศาสตร์
 3. นักเรียนไม่สามารถที่จะคำนวณตามที่โจทย์ต้องการได้
- เมื่อไม่สามารถทราบว่าการที่นักเรียนทำข้อสอบหรือแก้ปัญหาคณิตศาสตร์ข้อนั้น

ผิดเนื่องจากสาเหตุใด เพราะคำตอบผิดของนักเรียนเกิดจากความคลาดเคลื่อน 3 ประการ คือ ความคลาดเคลื่อนเนื่องมาจากความเข้าใจในการอ่าน ความคลาดเคลื่อนเนื่องมาจากการใช้กระบวนการหรือการกระทำทางคณิตศาสตร์ และความคลาดเคลื่อนเนื่องมาจากการคำนวณ (Forsyth and Spratt, 1980 : 31) ดังนั้นการที่จะหาว่านักเรียนทำข้อสอบวิชาคณิตศาสตร์ หรือแก้ ปัญหาโจทย์คณิตศาสตร์ไม่ได้ เป็นเพราะสาเหตุใดเราจึงจำเป็นต้องสร้างแบบทดสอบขึ้นเพื่อวัด สาเหตุนั้น ๆ ขึ้น ซึ่งลัวน สายยศ (2518 : 28) ได้เสนอแนะการสร้างข้อสอบเพื่อวัดความสามารถ หาเหตุผลในการแก้โจทย์ปัญหาทางคณิตศาสตร์ได้ว่า ข้อสอบคณิตศาสตร์ที่ใช้วัดความสามารถ ในการหาเหตุผลในการแก้โจทย์ปัญหานั้น ความเข้าใจด้านภาษามีอิทธิพลอยู่มาก เพราะข้อค ำถามต้องใช้วิธีอธิบายด้วยภาษาเป็นส่วนใหญ่ เมื่ออ่านโจทย์ไม่เข้าใจแล้ว โอกาสที่จะทำแบบ ทดสอบนั้นถูกต้องจะมีน้อย ดังนั้นผู้สร้างคำถามหรือโจทย์จะต้องพยายามใช้ภาษาที่คนทั่วไปเข้าใจ ง่าย และมีความเป็นปรนัยในตัวเอง และเนื่องจากจุดมุ่งหมายของข้อสอบต้องการที่จะวัด ว่าหลังจากนักเรียนได้เรียนรู้เรื่องใดเรื่องหนึ่งไปแล้ว นักเรียนมีความบกพร่องด้านการหาเหตุผล ทางคณิตศาสตร์ในเรื่องนั้นเพียงใด จะได้นำไปแก้ไขได้ถูกต้อง ทันทีที่ การแก้โจทย์ปัญหาทาง คณิตศาสตร์ได้นั้น ต้องอาศัยเหตุผลขั้นตอนความคิดที่ถูกต้อง การนำคณิตศาสตร์ไปใช้ประโยชน์ จึงจะมีโอกาสผิดพลาดได้น้อย สำหรับในเรื่องเกี่ยวกับการหาข้อบกพร่องในการแก้โจทย์ปัญหา ทางคณิตศาสตร์นี้ มีผู้ศึกษาวิจัยไว้คือ ฟอริท และสปรัท (Forsyth and Spratt, 1980 : 31-34) ได้เสนออาจสร้างข้อสอบที่วัดทักษะการหากระบวนการแก้ปัญหาคณิตศาสตร์ โดยใช้ข้อ สอบที่อ่านเข้าใจง่าย และใช้รูปแบบคำตอบเป็นแบบกระบวนการคิดของโจทย์คณิตศาสตร์ และ หลังจากนั้น ฟอริท และแอนสเลย์ (Forsyth and andley, 1982 : 257-253) ได้ศึกษาในเรื่องเดียวกัน แต่ให้นักเรียนหาคำตอบ โดยใช้เครื่องคิดเลขแทนการใช้ทักษะการคำ

จากการศึกษาการแก้ปัญหาคณิตศาสตร์ในครั้งนี้ ผู้วิจัยต้องการที่จะศึกษาแบบ ทดสอบ เมื่อลดจำนวนข้อให้มีจำนวนน้อยข้อนั้นควรจะมีจำนวนข้อเท่าใด จึงจะมีค่าความเชื่อมั่น สูงสุด เพราะแบบทดสอบวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหามีความยาวมากเกินไป อาจจะไม่ ทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงตามทฤษฎี เนื่องจากนักเรียนจะล้า และเกิดความเบื่อ หน่าย ที่จะคิดตอบคำถามต่าง ๆ ทำให้นักเรียนไม่ตั้งใจทำแบบทดสอบนั้น ๆ ทำให้เกิดการเดา คำตอบเพื่อให้การทดสอบเสร็จสิ้นโดยเร็ว ส่งผลให้การวิจัยผิดพลาดอย่างน่าเสียดาย ดังนั้น ผู้วิจัยจึงใช้ แบบทดสอบเลือกตอบวิชาคณิตศาสตร์ฉบับโจทย์ปัญหาชั้นประถมศึกษาปีที่ 5 เรื่อง โจทย์ปัญหาเศษส่วน โจทย์ปัญหาคณนิยมและโจทย์ปัญหาร้อยละ โดยใช้ข้อคำถามที่เกี่ยวกับ โจทย์ปัญหาที่มีภาษามาเกี่ยวข้อง แล้วให้นักเรียนคิดแก้ปัญหาเพื่อหาคำตอบจากข้อคำถามที่

กำหนดให้

ความเชื่อมั่นของแบบทดสอบ

นักการศึกษาและนักจิตวิทยาได้ให้ความหมายของ "ความเชื่อมั่น" ไว้ต่าง ๆ กันดังนี้
กรอนลันด์ (Gronlund, 1976:105) ให้ความหมายว่า เป็นความคงที่ของคะแนนในการทดสอบ หรือความคงที่จากการประเมินครั้งแรกและครั้งอื่น ๆ

อนาสตาซิ (Anastasi, 1968:105) ให้ความหมายว่า เป็นความคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันแต่ต่างเวลาและต่างโอกาสกัน

ลินด์วอลล์ และนิคโค (Lind Vall and Nitko, 1967:126) ให้ความหมายว่า ค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบสองครั้ง โดยใช้แบบทดสอบฉบับเดียวกัน สอบในเวลาที่แตกต่างกัน

ลอร์ด และโนวิก (Lord and Novick, 1967:46) ให้ความหมายว่าเป็นความคงที่ของคะแนนที่ได้จากการตอบแบบทดสอบซ้ำ และคะแนนที่ได้จากการตอบแบบทดสอบทั้งสองครั้ง เป็นอิสระไม่ขึ้นกับความคลาดเคลื่อนของการวัดใด ๆ

นูนนาลลี (Nunnally, 1964:59) กล่าวว่า ความเชื่อมั่นเป็นส่วนระหว่างความความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการทดสอบ

ซวาล แพร์ตกุล (2516:136) สำเริง บุญเรืองรัตน์ (2529: 3 5) บุญเชิด ภิญโญ อนันตพงษ์ (2521:269) ได้ให้ความหมายว่า เป็นความคงที่ของคะแนนที่ได้จากการทดสอบ แบบทดสอบที่มีความเชื่อมั่นสูง ไม่ว่าจะวัดบุคคลเดียวกันกี่ครั้งก็จะได้คะแนนเท่าเดิม ในทางตรงข้ามแบบทดสอบที่มีความเชื่อมั่นต่ำคะแนนที่ได้จากการสอบสองครั้งของบุคคลเดียวกัน จะมีการเปลี่ยนแปลงมาก

ทฤษฎีความเชื่อมั่น (Theory of Reliability)

ความเชื่อมั่นในทฤษฎีมาตรฐานเดิม (Reliability in Classical Theory) กล่าวว่า การวัดที่เชื่อถือได้ต้องให้ผลที่แม่นยำ แต่ผลจากการวัดล้วนแต่ขาดความแม่นยำเสมอไม่มากนักน้อยซึ่งเป็นไปตามข้อสมมติฐานของสเปียร์แมน (Spearman, 1910) ที่ว่า คะแนนที่ได้จากการสอบ (Observed Scores) จะประกอบด้วย คะแนนจริง (True Scores) และคะแนนความคลาดเคลื่อน (Error Scores) ดังสมการ

$$X = T + E$$

เมื่อ X = คะแนนที่ได้จากการทดสอบ

T = คะแนนจริง

E = คะแนนความคลาดเคลื่อน

คะแนนความคลาดเคลื่อนคือ คะแนนที่เกิดจากความคลาดเคลื่อนในการวัด แบ่งออกเป็น 2 ประเภท คือ ความคลาดเคลื่อนที่เกิดจากการสุ่ม (Random Error) กับความคลาดเคลื่อนอย่างมีระบบ (Systematic Error) สำหรับความคลาดเคลื่อนประเภทหลังจะไม่มีผลกระทบต่อความเชื่อมั่น ส่วนความคลาดเคลื่อนที่เกิดจากการสุ่มจะมีค่าทั้งบวกและลบจึงหักล้างกันหมด และเนื่องจากความคลาดเคลื่อนไม่มีความสัมพันธ์กับคะแนนจริง จึงเขียนสมการ ความแปรปรวนของแบบทดสอบได้ ดังนี้

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

เมื่อ σ_X^2 = ความแปรปรวนของคะแนนสอบ

σ_T^2 = ความแปรปรวนของคะแนนจริง

σ_E^2 = ความแปรปรวนของคะแนนความคลาดเคลื่อน

จากทฤษฎีความเชื่อมั่น ($\rho_{xx'}$) คืออัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สอบ ซึ่งเขียนเป็นสมการได้ดังนี้

$$\rho_{xx'} = \frac{\sigma_T^2}{\sigma_X^2}$$

ซึ่งสามารถเขียนได้อีกอย่างหนึ่งดังนี้

$$\rho_{xx'} = 1 - \frac{\sigma_E^2}{\sigma_X^2}$$

(Mehrens and Lehmann, 1984 : 268-269)

การตีความหมายสัมประสิทธิ์ความเชื่อมั่นแบบต่าง ๆ

แอลเลน และเยน (Allen and Yen, 1979 : 73-76) ได้เสนอการพิจารณาค่าสัมประสิทธิ์ความเชื่อมั่น ($\rho_{xx'}$) ไว้ 6 ประการดังนี้

1. $\rho_{xx'}$ คือสหสัมพันธ์ระหว่างคะแนนสอบที่ได้จากแบบทดสอบคู่ขนานของผู้สอบแต่ละคน ถ้าคะแนนสอบของแต่ละคนได้เหมือนกัน เมื่อใช้แบบทดสอบคู่ขนานแบบทดสอบก็มีความเชื่อมั่นที่สมมูลกับถ้าคะแนนจากการสอบที่ได้จากแบบทดสอบคู่ขนานไม่มีความสัมพันธ์กันแบบทดสอบก็ไม่มีคามเชื่อมั่น

2. $\rho_{xx'}^2$ เท่ากับสัดส่วนของความแปรปรวนใน x ที่อธิบายโดยความสัมพันธ์เชิงเส้นตรงกับ x เป็นการตีความหมายแบบมาตรฐานสำหรับสัมประสิทธิ์สหสัมพันธ์แบบ เพียร์สัน ซึ่งกำลังสองของสหสัมพันธ์สามารถตีความหมายได้ว่าเป็นสัดส่วนของความแปรปรวนในตัวแปรหนึ่งที่อธิบายได้ด้วยความสัมพันธ์เชิงเส้นตรงกับตัวแปรอื่น ๆ ดังนั้น $\rho_{xx'}^2$ เป็นสัดส่วนของความแปรปรวนในคะแนนในแบบทดสอบคู่ขนาน

3. $\rho_{xx'}$ คือ ความแปรปรวนของคะแนนจริง (σ_T^2) หารด้วยความแปรปรวนของคะแนนสอบ (σ_X^2) ถ้าค่าความแปรปรวนของคะแนนสอบ เท่ากับความแปรปรวนของคะแนนจริง แล้ว $\rho_{xx'} = 1$ และความแปรปรวนของความคลาดเคลื่อน (σ_E^2) เท่ากับ 0 จะสรุปได้ว่า การวัดปราศจากความคลาดเคลื่อน ถ้าความเชื่อมั่นน้อยกว่า 1 แสดงว่าการวัดมีความคลาดเคลื่อน ถ้าความแปรปรวนของคะแนนสอบเท่ากับค่าแปรปรวนของความคลาดเคลื่อน จะได้ $\rho_{xx'} = 0$ ดังนั้น ถ้าความแปรปรวนของความคลาดเคลื่อนมีน้อย คะแนนสอบจะมีค่าใกล้เคียงกับคะแนนจริง

4. $\rho_{xx'}$ คือ กำลังสองของสหสัมพันธ์ระหว่างคะแนนสอบกับคะแนนจริง ρ_{XT}^2 เช่น ถ้าความเชื่อมั่นมีค่าเท่ากับ .64 แล้วความสัมพันธ์ระหว่างคะแนนที่สอบกับคะแนนจริงมีค่าเท่ากับ .80 ซึ่งสรุปได้ว่า

$$\rho_{xx'} < \rho_{XT}$$

5. $\rho_{xx'} = 1 - \rho_{XE}^2$ หมายถึง สัมประสิทธิ์แห่งความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยกำลังสองของความสัมพันธ์ระหว่างคะแนนสอบกับคะแนนความคลาดเคลื่อน ถ้า $\rho_{XE} = 0$ จะได้ $\rho_{xx'} = 1$

6. $\rho_{xx'} = 1 - \frac{\sigma_E^2}{\sigma_X^2}$ หมายถึงค่าความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยอัตราส่วนระหว่างความแปรปรวนคลาดเคลื่อนกับความแปรปรวนของคะแนนสอบ

$$\text{ถ้า } \rho_{xx'} = 1 \text{ แล้ว } \sigma_E^2 = 0$$

$$\text{ถ้า } \rho_{xx'} = 0 \text{ แล้ว } \sigma_E^2 = \sigma_X^2$$

ดังนั้นความแปรปรวนของคะแนนสอบจะมีผลกระทบต่อความเชื่อมั่น จะเห็นว่าความเชื่อมั่นที่ประมาณได้จากกลุ่มวิวิธพันธ์จะมีแนวโน้มสูงกว่าการประมาณค่าจากกลุ่มเอกพันธ์

วิธีประมาณค่าความเชื่อมั่น

แอลเลน และเยน (Allen and Yen, 1979:88) กล่าวว่า วิธีประมาณค่าความเชื่อมั่นที่ต่างกันจะได้ค่าความเชื่อมั่นที่ต่างกัน สำหรับแบบทดสอบอาศัยความเร็ว ควรใช้แบบสอบซ้ำ หรือแบบทดสอบคู่ขนาน เพราะการหาค่าความเชื่อมั่นโดยใช้การแบ่งครึ่งแบบทดสอบอาจทำให้ค่ามากกว่าความจริง ส่วนการใช้สัมประสิทธิ์แอลฟา (α) และวิธีของ Kuder-Richardson จะให้ค่าที่ต่ำกว่าความเป็นจริง

เฟอร์กูสัน (Ferguson, 1966:365-366) และสแตนเลย์ และฮอปกินส์ (Standley and Hopkins, 1972:122-127) ได้กล่าวว่ามียุทธวิธีการหาค่าสัมประสิทธิ์ความเชื่อมั่น 4 วิธี ดังนี้

1. วิธีสอบซ้ำ หรือเรียกว่าสัมประสิทธิ์ของความคงที่ (Coefficient of stability) เป็นการนำแบบทดสอบฉบับเดียวกันไปทดสอบกับคน ๆ เดียวซ้ำสองครั้งในช่วงเวลาที่แตกต่างกันพอสมควร คะแนนที่ได้จากการทดสอบทั้งสองครั้งนี้มีความสัมพันธ์กัน ค่าสหสัมพันธ์ที่ได้จะเป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

2. วิธีใช้แบบทดสอบคู่ขนาน เป็นการนำแบบทดสอบที่มีลักษณะคู่ขนานกัน หรือเท่าเทียมกัน โดยมีเนื้อหา ค่าเฉลี่ย และความแปรปรวนเท่ากันไปทดสอบในเวลาเดียวกัน หรือในเวลาที่แตกต่างกัน คะแนนที่ได้จากแบบทดสอบทั้งสองฉบับมีสหสัมพันธ์และค่าสหสัมพันธ์ที่ได้คือค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

3. วิธีแบ่งครึ่งแบบทดสอบ เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มเดียว แล้วแบ่งครึ่งแบบทดสอบเป็นชุดคะแนนของข้อดีและข้อคู่หน้าคะแนนทั้ง 2 ชุดไปหาความสัมพันธ์จากนั้นปรับขยายด้วยสูตรของ สเปียร์แมน - บราวน์ ค่าที่ได้คือค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4. วิธีวัดความสอดคล้องภายในของแบบทดสอบ เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มหนึ่งและนำไปหาสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ โดยใช้สูตร KR-20

เมห์เรนส์ และเลห์แมนน์ (Mehrens and Lehmann, 1984:271-272) กล่าวถึงการประมาณค่าความเชื่อมั่น มีวิธีการ ดังนี้

1. วิธีแบบสอบซ้ำ

2. วิธีใช้แบบทดสอบคูขนาน
3. วิธีใช้แบบทดสอบคูขนานและสอบซ้ำ
4. วิธีวัดความสอดคล้องภายใน
 - 4.1 วิธีแบ่งครึ่งข้อสอบ
 - 4.2 วิธีของ Kuder-Richardson
 - 4.3 วิธีสัมประสิทธิ์แอลฟา (Coefficient Alpha)
 - 4.4 วิธีวิเคราะห์ความแปรปรวนของ ฮอยท์ (Hoyt's)
5. ความเชื่อมั่นของผู้ให้คะแนน

จากการศึกษาเกี่ยวกับความเชื่อมั่นในครั้งนี ทำให้ผู้วิจัยเกิดแนวคิดที่ว่า น่าจะมีวิธีการที่จะทำให้แบบทดสอบมีความเชื่อมั่นสูงขึ้นนั้น อาจทำได้โดยการลดจำนวนข้อสอบลง โดยการคัดเลือกแบบทดสอบให้มีความยากง่ายเหมาะสมกับนักเรียน มีความเป็นปรนัยในการให้คะแนน ในการศึกษาครั้งนี้ ผู้วิจัยจึงศึกษาเกี่ยวกับความยาวของแบบทดสอบเมื่อลดความยาวให้สั้นลง หรือให้มีจำนวนน้อยข้อนั้น ควรจะมีจำนวนข้อเท่าใดจึงจะมีค่าความเชื่อมั่นสูงสุด โดยศึกษาว่าแบบทดสอบเมื่อลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครูจะมีผลทำให้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันหรือไม่ สำหรับการคำนวณสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบนั้น ใช้สูตรของเฟลด์ต์ (Feldt) เนื่องจากเป็นแบบทดสอบ K ส่วน มีการทดสอบเพียงครั้งเดียว เนื้อหาของแบบทดสอบเป็นเอกพันธ์ ข้อสอบเป็นปรนัยชนิดเลือกตอบซึ่งให้คะแนน 0-1

การคัดเลือกข้อสอบ

จุดมุ่งหมายของการวิเคราะห์ข้อสอบเพื่อปรับปรุงคุณภาพของแบบทดสอบ ให้มีความเที่ยงตรง และความเชื่อมั่นสูงสุด ดังนั้นเกณฑ์ในการคัดเลือกข้อสอบก็คือ จะต้องเลือกข้อสอบที่จะทำให้แบบทดสอบมีความเที่ยงตรงและความเชื่อมั่นสูง (บุญเรียง ขจรศิลป์, 2533:117) แบบทดสอบที่ประกอบด้วยข้อสอบที่ง่ายมาก ๆ จะมีความเชื่อมั่นต่ำ แต่แบบทดสอบที่ประกอบด้วยข้อสอบที่มีความง่ายอยู่ระดับปานกลางคือ ไม่ง่ายจนเกินไป หรือยากจนเกินไปจะมีความเชื่อมั่นสูง ในขณะเดียวกันแบบทดสอบที่ประกอบด้วยข้อสอบที่มีค่าอำนาจจำแนกสูงเท่าใด ก็จะทำให้ค่าความเชื่อมั่นสูงมากขึ้นเท่านั้น ดังจะเห็นได้จากตาราง 1

ตาราง 1 แสดงความสัมพันธ์ระหว่าง อำนาจจำแนกของข้อสอบกับค่าความเชื่อมั่นของแบบทดสอบ ที่ประกอบด้วยข้อสอบ 100 ข้อ

ค่าเฉลี่ยของดัชนีอำนาจจำแนก	ความเชื่อมั่นของแบบทดสอบ
0.12	0.00
0.16	0.42
0.21	0.63
0.30	0.84
0.40	0.92
0.50	0.95

1. ความยากง่ายของข้อสอบ (Item difficulty)

ความหมายของ “ความยากง่ายของข้อสอบ” (Item difficulty) ความยากง่ายของข้อสอบสามารถให้ความหมายได้หลาย ๆ แนวทาง แต่โดยทั่ว ๆ ไปจะนิยามว่า “ความยากง่ายของข้อสอบคือ สัดส่วนหรือเปอร์เซ็นต์ของจำนวนผู้สอบที่ตอบข้อคำถามนั้นถูกต้อง” (Ebel, 1979:395) และเขียนแทนค่าความยากง่ายด้วย P ความยากง่ายของข้อสอบแต่ละข้อมีลักษณะดังนี้ 1.) มีค่าตั้งแต่ .00-1.00 2.) ข้อง่าย ค่าของ p จะมาก 3.) ข้อยากค่าของ p จะน้อย 4.) ข้อสอบที่มีผู้ตอบถูกหมดทุกคน แสดงว่าข้อสอบข้อนั้นง่ายมาก 5.) ข้อสอบที่ไม่มีผู้ตอบถูกเลย แสดงว่าข้อสอบข้อนั้นยากมาก 6.) ข้อสอบที่มีค่า $p = .50$ ถือว่าข้อสอบข้อนั้นยากพอเหมาะ หมายความว่า เป็นข้อสอบที่มี 50 เปอร์เซ็นต์ หรือครึ่งหนึ่งของผู้สอบตอบถูก อย่างไรก็ตามถ้าข้อสอบข้อ 1 และข้อ 5 ต่างก็มีค่า $p = .50$ ไม่ได้หมายความว่าผู้ที่ตอบข้อ 1 ถูก จะต้องตอบข้อ 5 ถูกด้วย ผู้ที่ตอบข้อ 5 ถูก อาจจะเป็นผู้สอบคนอื่นแต่มีจำนวนเท่ากับครึ่งหนึ่งของผู้สอบเช่นกัน 7.) ค่า P ยิ่งมาก ข้อสอบยิ่งง่าย ส่วนค่า P น้อย ๆ ข้อสอบจะยาก

2. ค่าอำนาจจำแนก (Discrimination)

ความหมายของ “ค่าอำนาจจำแนก” (Discrimination) อำนาจจำแนกของข้อสอบเป็นคุณสมบัติเฉพาะของข้อคำถาม ในการแบ่งเด็กออกตามระดับความสามารถคือกลุ่มเก่งกับกลุ่มอ่อนได้ โดยถือหลักง่าย ๆ ว่า ถ้าข้อสอบที่มีคนที่ได้คะแนนรวมมากซึ่งจัดอยู่ในกลุ่มเก่งควรตอบได้ถูก ส่วนคนที่ได้คะแนนรวมน้อยซึ่งจัดอยู่ในกลุ่มอ่อนควรตอบผิด ค่าอำนาจจำแนกนี้แทนค่า

- ถ้าวัดได้ตรงจุดประสงค์น้อย ควรตัดออกไป
- 2.) ข้อสอบวัดได้ครอบคลุมเนื้อหาที่จะทดสอบหรือไม่
 - ถ้าวัดได้ตรงเนื้อหาที่สุด ควรเก็บข้อนั้นไว้
 - ถ้าวัดได้ตรงเนื้อหาบ่อย ควรตัดออกไป

โดยนำมาทำเป็นแบบประเมินความสอดคล้องของข้อสอบกับจุดประสงค์และเนื้อหาตามวิธีของ ไรวีนลลี และแฮมเบิลตัน (บุญเชิด ภิญญอนันตพงษ์, 2527: 68-69 ; อ้างอิงมาจาก Hambleton et al. 1978:34) หลังจากครูผู้สอนกำหนดคะแนนของการพิจารณาความสอดคล้องระหว่างข้อคำถามกับเกณฑ์การพิจารณาแล้ว นำเอาคะแนนรวมของข้อสอบแต่ละข้อจากครูผู้สอนทั้งหมดไปหาค่าสัดส่วนของจำนวนผู้ตัดสินข้อนั้นกับจำนวนผู้พิจารณาทั้งหมด เรียกว่า ค่าดัชนีความสอดคล้อง (IOC) โดยพิจารณาคัดเลือกข้อที่มีค่าดัชนีความสอดคล้องอยู่ระหว่าง 0.5 - 1

ความยาวของแบบทดสอบ

ความยาวของแบบทดสอบ ก็คือจำนวนข้อของแบบทดสอบ กรอนลันด์ (Grolund, 1976 :117) กล่าวไว้ว่า การเพิ่มจำนวนข้อสอบจะทำให้ค่าความเชื่อมั่นสูงขึ้น ซึ่งเป็นไปตามสูตรของ สเปียร์แมน – บราวน์ นอกจากนั้น แมกนัสสัน (Magnusson, 1967:68-74) ได้กล่าวว่า การเพิ่มความยาวของแบบทดสอบด้วยการรวมแบบทดสอบที่ขนานกันหลายๆฟอร์มเข้าด้วยกันทำให้ค่าความคลาดเคลื่อนต่าง ๆ ถูกขจัดออกไปหมดจนเหลือแต่คะแนนจริงซึ่งจะทำให้ค่าความเชื่อมั่นเพิ่มขึ้นนั่นเอง

ดังนั้น เมื่อรวมแบบทดสอบที่ขนานกันเข้าเป็นแบบทดสอบฉบับเดียว จำนวนมาก ๆ จนความยาวเป็นอนันต์ (Infinity) แล้วค่าความเชื่อมั่นจะมีค่าเป็น 1.00 อย่างไรก็ตามการเพิ่มจำนวนข้อสอบ เพื่อให้ค่าความเชื่อมั่นสูงขึ้นนั้น จะต้องขึ้นอยู่กับจำนวนข้อที่เพิ่มขึ้นจะต้องมีคุณภาพเหมือนข้อสอบเดิมและการเพิ่มนั้นจะต้องให้เหมาะสมกับเวลาด้วย อัดัมส์ (Adam, 1964 : 97) ส่วนอนันต์ ศรีโสภา (2525:170-171) กล่าวว่า การที่จะทำให้แบบทดสอบมีความเชื่อมั่นสูงขึ้น อาจทำได้โดยการเพิ่มจำนวนข้อสอบ แต่ถ้าเพิ่มข้อสอบแล้วไม่ทำให้ค่าความเชื่อมั่นสูงขึ้นอาจมาจากสาเหตุสามประการ คือ

1. ข้อสอบต่าง ๆ จาก Domain มีสหสัมพันธ์ระหว่างกันต่ำมาก
2. ข้อสอบเกาะกันเป็นกลุ่ม ๆ (Clusters) คือแต่ละข้อภายในกลุ่มมีสหสัมพันธ์กันสูง แต่มีสหสัมพันธ์กับข้อในกลุ่มอื่นต่ำ ทำให้สหสัมพันธ์กับคะแนนรวมต่ำ
3. ข้อสอบบางข้อมีสหสัมพันธ์อย่างสูงกับบางข้อ แต่มีสหสัมพันธ์กับข้ออื่น ๆ เกือบเท่ากับศูนย์

ในทางปฏิบัติจะเห็นว่ากรณีที่เรเพิ่มข้อสอบเข้าไปไม่ทำให้ค่าความเชื่อมั่นสูงมักมีสาเหตุหลายประการ เช่น ข้อสอบที่เพิ่มเข้าไปต้องมีค่าความยากและค่าอำนาจจำแนกเหมือนข้อสอบเดิมซึ่งเป็นการยากมากที่เราจะสร้างข้อสอบให้มีค่าความยาก และอำนาจจำแนกเท่ากัน นอกจากนั้นความลำเอียงของข้อมูลของกลุ่มตัวอย่างก็เป็นตัวแปรตัวหนึ่งที่ทำให้ความเชื่อมั่นของแบบทดสอบต่ำ จากปัญหาดังกล่าวผู้วิจัยจึงต้องการศึกษาถึงความยาวของแบบทดสอบที่มีผลต่อค่าความเชื่อมั่นโดยการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ และใช้เกณฑ์ดุลพินิจของครู

งานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า

วิเชียร เกตุสิงห์ (2523:240-243) ได้ทำการทดลองเพื่อลดจำนวนข้อในแบบทดสอบให้น้อยลง โดยนำแบบทดสอบวัดความถนัดทางการเรียน 5 ฉบับคือ ภาษา (คำศัพท์) คณิตศาสตร์ เหตุผล (ภาษา) เหตุผล (รูป) มิติสัมพันธ์ รวมจำนวนข้อสอบเกือบ 200 ข้อ ใช้เวลานานกว่า 2 ชั่วโมง (แบบทดสอบทั้ง 5 ฉบับได้ผ่านการวิเคราะห์รายข้อ และผ่านการคัดเลือกมาแล้ว) มาทำการวิเคราะห์องค์ประกอบเป็นรายฉบับ แล้วนำผลการวิเคราะห์นั้น มาพิจารณาเลือกข้อสอบด้วยวิธีต่าง ๆ กัน คือ

วิธีที่ 1 พิจารณาในแต่ละองค์ประกอบ มีข้อสอบใดบ้างที่มีค่าน้ำหนักองค์ประกอบ (Factor Loading) สูงตั้งแต่ .50 ขึ้นไป แล้วเลือกข้อเหล่านั้นมาทุกข้อ

วิธีที่ 2 ในแต่ละองค์ประกอบ เลือกเฉพาะข้อที่มีน้ำหนักองค์ประกอบสูงสุดมาเพียงข้อเดียวเพื่อเป็นตัวแทนขององค์ประกอบนั้น ๆ

วิธีที่ 3 ในแต่ละองค์ประกอบ เลือกข้อสอบที่มีค่าน้ำหนักองค์ประกอบต่ำกว่า .50 มาทุกข้อ ส่วนข้อที่มีค่าน้ำหนักองค์ประกอบสูงกว่า .50 เลือกที่มีค่าสูงสุดมาเพียงข้อเดียว

นำมาตรวจให้คะแนนใหม่แล้วคำนวณค่าสถิติพื้นฐาน ความเชื่อมั่น และหาค่าสหสัมพันธ์ระหว่างวิธีต่าง ๆ ทั้ง 3 วิธีกับคะแนนรวมทั้งฉบับที่ไม่ได้เลือกบางข้อออก จากการทดลองพบว่าวิธี ที่ 1 และวิธีที่ 3 ให้ผลใกล้เคียงกันมาก ถ้าพิจารณาในแง่ของความสัมพันธ์กับคะแนนรวมทั้งฉบับโดยไม่มีการเลือกข้อสอบออก ส่วนค่าความเชื่อมั่น บางฉบับใกล้เคียงกันทั้งสองวิธีคือ ฉบับภาษาและเหตุผล (ภาษา) นอกนั้นให้ค่าความเชื่อมั่นแตกต่างกัน คือฉบับที่มีข้อมากจะมีค่าความเชื่อมั่นสูงกว่าฉบับที่มีจำนวนข้อน้อย ส่วนสัมประสิทธิ์ของความแปรผัน (C.V.) วิธีที่ 1 ให้ผลใกล้เคียงกับการใช้ข้อสอบทุกข้อมากที่สุด

จากการทดลอง ถ้ามีความจำเป็นจะต้องใช้ข้อสอบน้อยข้อเพื่อประหยัดทั้งเวลา ค่าใช้จ่ายวิธีการนี้ก็จะเป็นวิธีที่จะใช้ได้เพราะสามารถลดข้อสอบจากเดิมเหลือ 65 % ของข้อสอบที่มี

คุณภาพที่มีอยู่ทั้งหมด

ซิมเมอร์แมน และฮัมฟรี (วารุณี ปิครวัชชัย, 2513:9; อ้างอิงมาจาก Zimmermann and Humphrey, 1953:460-461) ได้ศึกษาถึงค่าความเชื่อมั่นของแบบทดสอบเลือกตอบพบว่า การกำจัดตัวลองที่มีค่าอำนาจจำแนกต่ำสุดออกจากตัวเลือกของแบบทดสอบที่มี 5 ตัวเลือก จะทำให้ค่าความเชื่อมั่นของแบบทดสอบสูงขึ้น และทำให้เวลาในการตอบแต่ละข้อลดลง

บูเดสคู และเนโว (Budescu and Nevo, 1985:183-196) ได้ศึกษาตัวเลือกที่เหมาะสมเพื่อหาข้อสันนิษฐานเกี่ยวกับสัดส่วน โดยใช้กลุ่มตัวอย่าง 1018 คน ที่สอบเข้ามหาวิทยาลัยไฮฟา (The University of Haifa) เครื่องมือในการวิจัยเป็นแบบทดสอบ 3 ฉบับ ประกอบด้วย แบบทดสอบคำศัพท์ 60 ข้อ แบบทดสอบความเข้าใจในการอ่าน 24 ข้อ แบบทดสอบเหตุผลทางคณิตศาสตร์ 25 ข้อ ครั้งแรกใช้แบบทดสอบ 5 ตัวเลือก แล้วนำมาแบ่งเป็นแบบทดสอบสภาพการณ์ต่าง ๆ 6 แบบ คือ แบบทดสอบที่มี 4, 3 และ 2 ตัวเลือก โดยการสุ่มตัวลองตัวสุดท้ายออก และแบบทดสอบที่มี 4, 3 และ 2 ตัวเลือก โดยการคัดตัวลองที่มีค่าอำนาจจำแนกต่ำสุดออก ผลการศึกษาพบว่า แบบทดสอบทั้ง 3 ฉบับที่มีจำนวนตัวเลือกมากจะให้ค่าความแปรปรวนมาก ค่าความเชื่อมั่นและค่าอำนาจจำแนกสูงกว่าแบบทดสอบที่มีจำนวนตัวเลือกน้อย

ผดุงชาติ สุวรรณวงศ์ (2516:42-43) ได้ทำการวิจัยเรื่องความเชื่อมั่นและความเที่ยงตรงของข้อสอบที่เพิ่มความยาว โดยให้ความคิดว่า ถ้าขยายข้อสอบให้ยาวขึ้นความเชื่อมั่นจะสูงขึ้น ผลการวิจัยพบว่ามีข้อสอบบางฉบับเมื่อเพิ่มความยาวขึ้นแล้วมีค่าความเชื่อมั่นต่ำลงทั้งนี้เป็นเพราะข้อสอบที่เพิ่มบางข้อมีความสัมพันธ์ทางลบกับข้อเดิมจึงทำให้สหสัมพันธ์เฉลี่ยลดลง

ดร.น. หาญตระกูล (2519:1-43) ได้ทำการศึกษาเกี่ยวกับคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 7 เป็นแบบทดสอบเลือกตอบชุดเดียวกัน แต่มี 5, 4, 3 และ 2 ตัวเลือก โดยใช้การคัดตัวลองที่มีค่าอำนาจจำแนกต่ำสุดออกทีละตัวจนเหลือ 2 ตัวเลือก ผลการศึกษาพบว่า แบบทดสอบวัดทักษะทางคณิตศาสตร์ที่มี 4 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบ 2, 3 และ 5 ตัวเลือกอย่างมีนัยสำคัญทางสถิติ

บุญเรือน จรัสวิมล (2521:40-41) ได้ศึกษาการเพิ่มข้อสอบที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ โดยเพิ่มจำนวนข้อสอบในแบบทดสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ จากเดิม 30 ข้อ เป็น 60 ข้อ ข้อสอบที่เพิ่มในแต่ละฉบับมีช่วงระดับความยากต่างกันคือ .40 - .60, .30 - .70, และ .20 - .80 ผลการศึกษาพบว่า การเพิ่มจำนวนข้อสอบด้วยข้อสอบที่มีความยาวต่างกันไม่ทำให้ความเชื่อมั่นต่างกัน เมื่อนำแบบทดสอบที่เพิ่มความยาวไปเปรียบเทียบกับค่า ความเชื่อมั่นกับฉบับเดิมพบว่า แบบทดสอบที่เพิ่มความยาวที่มีช่วงระดับความยาก .30 - .70, และ .20 - .80 มีความเชื่อ

มันสูงกว่าฉบับเดิมอย่างมีนัยสำคัญทางสถิติที่ .01 และ .05 ตามลำดับ ส่วนแบบทดสอบที่เพิ่มจำนวนข้อในช่วงระดับความยาก .40-.60 ค่าความเชื่อมั่นสูงกว่าฉบับเดิมอย่างไม่มีนัยสำคัญทางสถิติ

ชนะ ทานะวงศ์ (2521:23-27) ได้ศึกษาเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนของแบบทดสอบปรนัยชนิดเลือกตอบที่มีจำนวนข้อ และตัวเลือกต่างกัน ในวิชาวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 4 โดยใช้แบบทดสอบวัดความเข้าใจทางวิทยาศาสตร์วัดทักษะทางวิทยาศาสตร์และวัดปฏิบัติทางวิทยาศาสตร์ ในแบบทดสอบหนึ่ง ๆ สร้างเป็น 4 ฉบับย่อย ๆ คือ แบบทดสอบ ชนิด 30 ข้อ 2 ตัวเลือก, 20 ข้อ 3 ตัวเลือก, 15 ข้อ 4 ตัวเลือก และ 12 ข้อ 5 ตัวเลือก ผลการศึกษาปรากฏว่าแบบทดสอบ 30 ข้อ 2 ตัวเลือกมีแนวโน้มให้ค่าความเชื่อมั่นสูงกว่าแบบทดสอบชนิดอื่น

ยศอารีย์ รวยธนพานิช (2534 : 56-58) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบวิชาคณิตศาสตร์ที่มีความยากแตกต่างกันคือ .40-.60, .30-.70 และ .20-.80 และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวนข้อต่างกันคือ 30,40 และ 50 ข้อ ผลการศึกษาพบว่า แบบทดสอบที่มีช่วงความยากต่างกันเมื่อมีจำนวนข้อเท่ากันมีค่าความเชื่อมั่นไม่ต่างกัน และแบบทดสอบที่มีความยาวต่างกันเมื่อมีช่วงความยากเท่ากัน มีค่าความเชื่อมั่นไม่ต่างกัน ยกเว้นที่มีช่วงความยาก .40-.60 แบบทดสอบที่มี 50,40 ข้อ มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 30 ข้อ ส่วนแบบทดสอบที่มี 50 และ 40 ข้อ มีค่าความเชื่อมั่นไม่ต่างกัน

อรนุช ธิติรักษ์พานิชย์ (2538 : 91-94) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดบุคคลิกภาพเยาวชนไทย เมื่อลดความยาวให้สั้นลง โดยมีความยาว 300 ข้อ แล้วลดความยาวลงเหลือ 210, 180, 150 และ 120 ข้อ นั้นความยาวของแบบทดสอบควรเป็นเท่าใดจึงจะเหมาะสมหรือยังมีค่าความเชื่อมั่นไม่แตกต่างไปจากฉบับเดิม ผลการศึกษาพบว่าแบบทดสอบที่ลดความยาวลงเหลือ 120 ข้อ มีค่าความเชื่อมั่นแตกต่างไปจากเดิม และความยาวของแบบทดสอบวัดบุคคลิกภาพเยาวชนไทยที่เหมาะสมคือ 150 ข้อ

จากการศึกษางานวิจัยที่เกี่ยวข้องกับความยาวของแบบทดสอบ ผู้วิจัยจึงเกิดแนวคิดว่าจะลดความยาวของแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหาค่าความเชื่อมั่นของแบบทดสอบจะแตกต่างกันหรือไม่ ด้วยเหตุผลที่ว่า แบบทดสอบวิชาคณิตศาสตร์ฉบับโจทย์ปัญหา ถ้ามีความยาวมากอาจทำให้นักเรียนเกิดความล้าความเบื่อหน่าย หรือไม่ตั้งใจในการที่จะคิดหาคำตอบ ทำให้เกิดการเดาคำตอบ เพื่อที่จะทำข้อสอบให้เสร็จโดยเร็ว ดังนั้นผู้วิจัยต้องการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ กับ

ใช้เกณฑ์ดุลพินิจของครู

สมมติฐานในการศึกษาค้นคว้า

1. แบบทดสอบที่ใช้เกณฑ์ค่าสถิติลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ มีค่าความเชื่อมั่นแตกต่างจากแบบทดสอบฉบับ 40 ข้อ
2. แบบทดสอบที่ใช้เกณฑ์ดุลพินิจของครูลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ มีค่าความเชื่อมั่นแตกต่างจากแบบทดสอบฉบับ 40 ข้อ
3. แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากันมีค่าความเชื่อมั่นแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดสุพรรณบุรี

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษา จังหวัดสุพรรณบุรี ซึ่งเลือกมาโดยวิธีการสุ่มแบบหลายขั้นตอน (Multistage Random Sampling) โดยมีลำดับขั้นตอนในการสุ่มตัวอย่างดังนี้

ขั้นที่ 1 แบ่งโรงเรียนทั้งหมดออกเป็น 10 อำเภอ แล้วสุ่มมา 3 อำเภอ ได้โรงเรียนที่เป็นกลุ่มตัวอย่างคือ อำเภอเมืองสุพรรณบุรี อำเภอบางปลาม้าและ อำเภอเดิมบางนางบวช

ขั้นที่ 2 สุ่มโรงเรียนในแต่ละอำเภอด้วยการสุ่มอย่างง่าย (Simple Random Sampling) แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม ดังนี้

1. กลุ่มตัวอย่างที่ใช้ในการสร้างและตรวจสอบคุณภาพเครื่องมือ จำนวน 4 โรงเรียน จำนวนนักเรียน 134 คน

2. กลุ่มตัวอย่างที่ใช้ในการวิจัย จำนวน 11 โรงเรียน จำนวนนักเรียน 608 คน

ตาราง 2 แสดงชื่อโรงเรียนที่เป็นกลุ่มตัวอย่าง ที่ใช้ในการสร้างเครื่องมือและใช้ในการ วิจัย

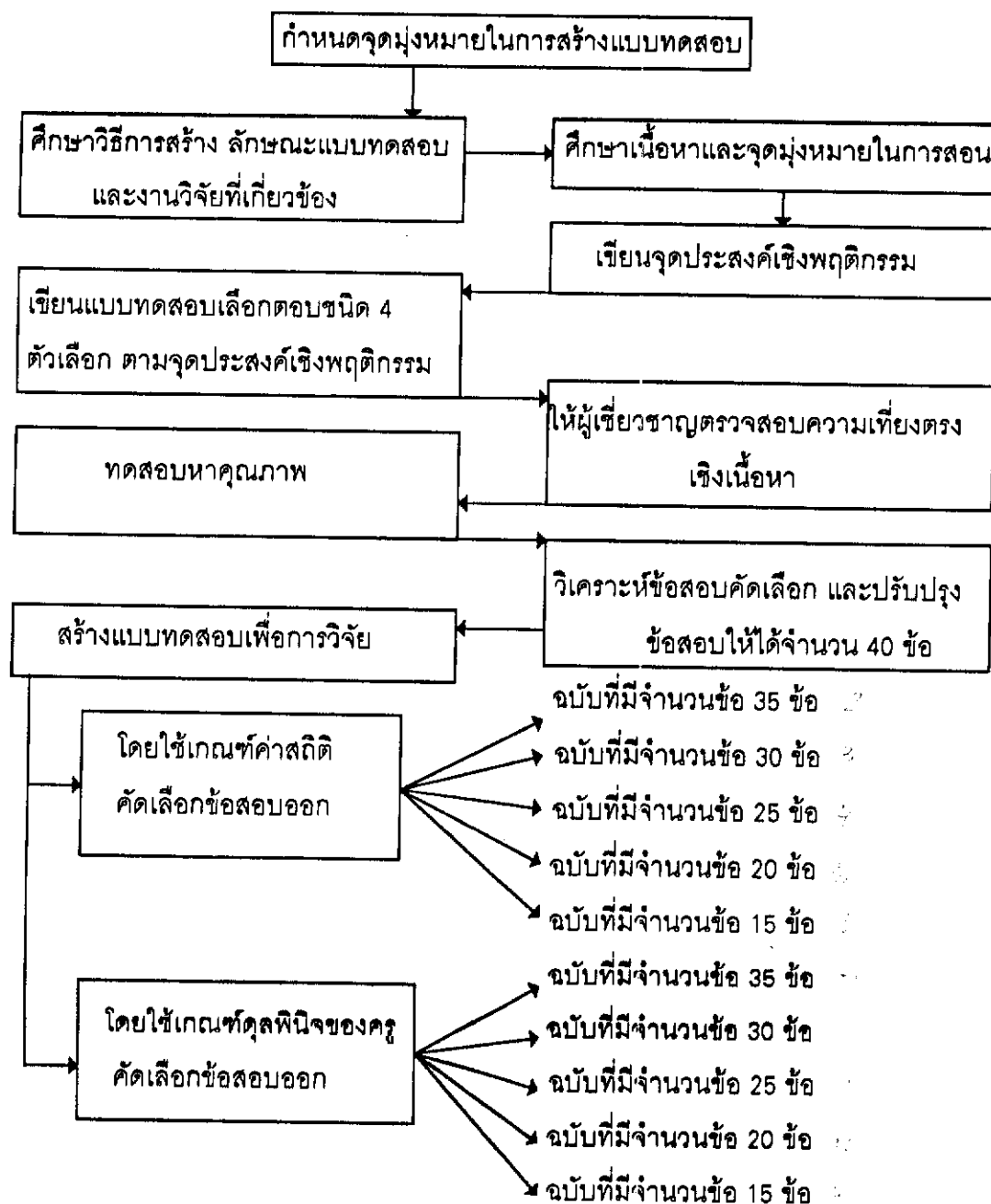
โรงเรียน	จำนวนนักเรียนที่ใช้ในการศึกษา	
	การสร้างเครื่องมือ	การวิจัย
สปอ.เมืองสุพรรณบุรี		
- โรงเรียนอนุบาลสุพรรณบุรี		150
- โรงเรียนวัดป่าเลไลย์		53
- โรงเรียนสุพรรณภูมิ	46	93
- โรงเรียนวัดพระธาตุ		61
สปอ.บางปลาม้า		
- โรงเรียนวัดสวนหงษ์	22	
- โรงเรียนบางปลาม้า		17
- โรงเรียนวัดดงครก	21	22
- โรงเรียนวัดช่องลม		49
- โรงเรียนวัดป่าพุก ษ์		39
สปอ.เดิมบางนางบวช		
- โรงเรียนวัดท่าช้าง		51
- โรงเรียนวัดเขาพระ	45	
- โรงเรียนวัดเดิมบางนางบวช		43
- โรงเรียนชุมชนวัดนางบวช		30
รวม	134	608

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ด้านโจทย์ปัญหาชั้นประถมศึกษาปีที่ 5 จำนวน 11 ฉบับซึ่งผู้วิจัยได้สร้างขึ้นให้ครอบคลุมเนื้อหาเรื่อง โจทย์ปัญหาเศษส่วน โจทย์ปัญหาคณิตนิยม และโจทย์ปัญหาร้อยละ โดยมีจำนวนข้อสอบในแบบทดสอบแต่ละฉบับที่ลดลงจากฉบับ 40 ข้อ เป็นฉบับ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ในการคัดเลือกข้อสอบออก

วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการวิจัย

ในการสร้างเครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้ดำเนินการสร้างแบบทดสอบเป็นลำดับขั้นตอน ดังนี้



ภาพประกอบ 1 แสดงลำดับขั้นตอนในการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

ในการสร้างเครื่องมือสำหรับการศึกษาค้นคว้าครั้งนี้ ได้ดำเนินการสร้างตามขั้นตอน ดังภาพประกอบ 1. ซึ่งมีรายละเอียดดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อสร้างแบบทดสอบเลือกตอบ วัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหาชั้นประถมศึกษาปีที่ 5 เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวนข้อต่างกัน

2. ศึกษาเนื้อหาและจุดมุ่งหมายในการสอนจากคู่มือครู การสอนคณิตศาสตร์ เรื่อง โจทย์ปัญหาเศษส่วน โจทย์ปัญหาคณิตศาสตร์ และโจทย์ปัญหาร้อยละ ชั้นประถมศึกษาปีที่ 5 ตามหลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง 2533)

3. จากผลการศึกษาในข้อ 2 ผู้วิจัยจึงนำมาเขียนเป็นจุดประสงค์เชิงพฤติกรรมดังนี้

1) เมื่อกำหนดโจทย์ปัญหาเศษส่วนหรือจำนวนคละ ที่มีตัวส่วนเท่ากันให้สองจำนวนสามารถหาผลลัพธ์โดยวิธีการบวก การลบได้

2) เมื่อกำหนดโจทย์ปัญหาเศษส่วนหรือจำนวนคละ ที่มีตัวส่วนไม่เท่ากันให้สองจำนวนสามารถหาผลลัพธ์โดยวิธีการบวก การลบได้

3) เมื่อกำหนดโจทย์ปัญหาเศษส่วนหรือจำนวนคละกับจำนวนเต็ม สามารถหาผลลัพธ์โดยวิธีการคูณ การหารได้

4) เมื่อกำหนดโจทย์ปัญหาคณิตศาสตร์ไม่เกิน 2 ตำแหน่ง สามารถหาผลลัพธ์โดยวิธี การบวก การลบได้

5) เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับร้อยละ รวมทั้งการซื้อขาย กำไร และขาดทุน สามารถหาผลลัพธ์ได้

4. นำจุดประสงค์เชิงพฤติกรรมที่วิเคราะห์แล้วมาสร้างเป็นแบบทดสอบเลือกตอบให้ได้จุดประสงค์ละ 12 ข้อ จะได้ข้อสอบทั้งหมด 60 ข้อ

5. นำแบบทดสอบเลือกตอบทั้ง 60 ข้อ ไปให้ผู้เชี่ยวชาญจำนวน 5 คน เป็นผู้พิจารณาตรวจสอบแบบทดสอบเป็นรายข้อว่าวัดตามจุดประสงค์เชิงพฤติกรรมหรือไม่ เกณฑ์ในการพิจารณาถ้าผู้เชี่ยวชาญ 3 ใน 5 คน เห็นว่าข้อสอบนั้นวัดตรงจุดประสงค์เชิงพฤติกรรมถือว่าข้อสอบนั้นใช้ได้ ผลการพิจารณาปรากฏว่า ข้อสอบวัดตรงจุดประสงค์ทั้ง 60 ข้อ ซึ่งข้อสอบ 60 ข้อนี้ ได้แสดงในภาคผนวก ก

6. นำแบบทดสอบที่ผู้เชี่ยวชาญพิจารณาแล้วไปทดสอบกับนักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 134 คน แล้วนำผลการสอบมาตรวจให้คะแนน ถ้าตอบถูกได้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบหรือตอบมากกว่า 1 คำตอบไม่ให้คะแนน

7. นำผลไปวิเคราะห์หาค่าความยาก และค่าอำนาจจำแนกของแบบทดสอบ โดยใช้เทคนิค 27 เปอร์เซนต์ และตารางสำเร็จรูปของ จุง เต ฟาน (Chung-Teh-Fan) คัดเลือกข้อสอบที่มีค่าความยากและค่าอำนาจจำแนกตามเกณฑ์การคัดเลือกที่ดี คือมีค่าความยากตั้งแต่ 0.20 - 0.80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป ให้ได้จำนวนข้อสอบจุดประสงค์ละ 8 ข้อรวมทั้งหมด 40 ข้อ

8. นำแบบทดสอบที่คัดเลือกข้อสอบไว้ 40 ข้อ มาพิจารณาคัดเลือกข้อสอบออก โดยใช้เกณฑ์ค่าสถิติดังนี้

8.1 พิจารณาข้อสอบที่มีค่าความยากและค่าอำนาจจำแนกที่มีค่าน้อยสุดตามลำดับออกไปจุดประสงค์ละ 1 ข้อ ได้ข้อสอบ 35 ข้อ

8.2 พิจารณาข้อสอบฉบับ 35 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 8.1 ได้ข้อสอบ 30 ข้อ

8.3 พิจารณาข้อสอบฉบับ 30 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 8.1 ได้ข้อสอบ 25 ข้อ

8.4 พิจารณาข้อสอบฉบับ 25 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 8.1 ได้ข้อสอบ 20 ข้อ

8.5 พิจารณาข้อสอบฉบับ 20 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 8.1 ได้ข้อสอบ 15 ข้อ

9. นำแบบทดสอบที่คัดเลือกข้อสอบไว้ 40 ข้อ มาพิจารณาคัดเลือกข้อสอบออก โดยใช้เกณฑ์ดุลพินิจของครูดังนี้

9.1 พิจารณาข้อสอบข้อที่มีค่าดัชนีความสอดคล้องต่ำที่สุดออกจุดประสงค์ละ 1 ข้อ ได้ข้อสอบ 35 ข้อ

9.2 พิจารณาข้อสอบฉบับ 35 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 9.1 ได้ข้อสอบ 30 ข้อ

9.3 พิจารณาข้อสอบฉบับ 30 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 9.1 ได้ข้อสอบ 25 ข้อ

9.4 พิจารณาข้อสอบฉบับ 25 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 9.1 ได้ข้อสอบ 20 ข้อ

9.5 พิจารณาข้อสอบฉบับ 20 ข้อ โดยใช้เกณฑ์การพิจารณาเกณฑ์เดียวกับข้อ 9.1 ได้ข้อสอบ 15 ข้อ

10. นำแบบทดสอบจำนวน 11 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างจำนวน 608 คน โดยแบ่งนักเรียนในแต่ละห้องเรียนออกเป็น 2-4 กลุ่ม ๆ ละเท่ากัน ให้นักเรียนแต่ละกลุ่มทำแบบทดสอบคนละ 1 ฉบับ เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ

ตัวอย่างข้อสอบ

จุดประสงค์เชิงพฤติกรรม

เมื่อกำหนดโจทย์ปัญหาทศนิยมไม่เกิน 2 ตำแหน่งให้สามารถหามลล์พรีโดยวิธีการบวกได้

ข้อสอบ

คุณพ่อเติมน้ำมันรถยนต์เมื่อตอนเช้า 22.5 ลิตร ตอนเย็นเติมอีก 37.5 ลิตร วันนี้คุณพ่อเติมน้ำมันทั้งหมดกี่ลิตร

- ก. 58.5 ลิตร
- ข. 59.0 ลิตร
- ค. 60.0 ลิตร
- ง. 61.5 ลิตร

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการตามขั้นตอน ดังนี้

1. นำแบบทดสอบไปดำเนินการทดสอบ ตามโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างที่ได้ขออนุญาตไว้
2. ดำเนินการสอบ โดยอธิบายให้ผู้สอบทุกคนเข้าใจถึงวัตถุประสงค์ และผลประโยชน์ที่ได้รับจากการตอบแบบทดสอบ
3. นำกระดาษคำตอบมาตรวจให้คะแนน
4. นำคะแนนมาวิเคราะห์ทางสถิติ และทดสอบตามสมมติฐานที่ตั้งไว้

ลำดับขั้นการวิเคราะห์ข้อมูล

ผู้วิจัยได้วิเคราะห์ข้อมูลตามลำดับขั้น ดังนี้

1. หาค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ
2. หาค่าความเชื่อมั่นของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ
3. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบตามสมมติฐานที่ตั้งไว้

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐาน คือ
 - 1.1 ค่าเฉลี่ย
 - 1.2 ค่าเฉลี่ยร้อยละ
 - 1.3 ส่วนเบี่ยงเบนมาตรฐาน
 - 1.4 ค่าสัมประสิทธิ์การกระจาย
2. สถิติวิเคราะห์ข้อสอบเป็นรายข้อ โดยใช้เทคนิค 27 เปอร์เซนต์ และตารางสำเร็จรูปของ จุง เต ฟาน (Chung-Teh-Fan) เพื่อหาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r)
3. ค่าความเชื่อมั่นของแบบทดสอบ คำนวณได้จากสูตรของเฟลด์ต์ (Feldt, 1975) (บุญเชิด ภิญโญนนตพงษ์. 2533:30)

$$r_{xx'} = \frac{1}{[1 - \sum \lambda_i^2]} \times \frac{[S_x^2 - \sum S_i^2]}{S_x^2}$$

เมื่อ

$$\lambda_i = \frac{S_{x_i x_1} + S_{x_i x_2} + \dots + S_{x_i x_k}}{S_x^2}$$

- เมื่อ
- $r_{xx'}$ = ค่าสัมประสิทธิ์ของความเชื่อมั่น
 - S_i^2 = ความแปรปรวนของคะแนนสอบแต่ละฉบับย่อย
 - S_x^2 = ความแปรปรวนของคะแนนทั้งฉบับ
 - $S_{x_i x_k}$ = ความแปรปรวนร่วมของคะแนนสอบแต่ละส่วนย่อย
 - λ_i = สัดส่วนของจำนวนข้อสอบแต่ละส่วนย่อย

4. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบ ทั้ง 11 ฉบับ โดยใช้สูตรของ Marascuilo (A. Ralph Hakstian and Thomas E. Whalen. 1976 : 224; อ้างอิงมาจาก Marascuilo. 1966)

เมื่อ n_k ไม่เท่ากันและ J_k ไม่เท่ากัน

$$M = \sum_{k=1}^k A_k - \frac{\left[\sum_{k=1}^k A_k (1 - r_{\alpha k})^{-\frac{1}{3}} \right]^{-2}}{\sum_{k=1}^k A_k (1 - r_{\alpha k})^{-\frac{2}{3}}}$$

$$\text{เมื่อ } A_k = \frac{(J_k - 1)(9n_k - 11)^2}{18J_k(n_k - 1)}$$

n_k = จำนวนคนในกลุ่มที่ k

J_k = จำนวนข้อในแบบทดสอบฉบับที่ k

$r_{\alpha k}$ = ค่าความเชื่อมั่นของแบบทดสอบฉบับที่ k

5. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบ เป็นรายคู่โดยใช้สูตรของ Scheffe' (A. Ralph Hakstian and Thomas E. Whalen, 1976 : 223-225; อ้างอิงมาจาก Marascuilo, 1966 และ Scheffe's 1959:67)

$$\sqrt{F} = \frac{P\hat{S}_i}{\sqrt{\text{Var}(\hat{\Psi})}} = \sqrt{1 - \alpha \chi_{k-1}^2}$$

$$\text{เมื่อ } P\hat{S}_i = \sum_{k=1}^k a_k \hat{\theta}_k$$

$$\text{และ } \text{Var}(\hat{\Psi}) = \sum_{k=1}^k a_k^2 \text{Var}(\hat{\theta}_k)$$

$$\hat{\theta}_k = (1 - r_{\alpha k})^{\frac{1}{3}}$$

$$\text{Var}(\hat{\theta}_k) = \frac{18J_k(n_k - 1)(1 - r_{\alpha k})^{\frac{2}{3}}}{(J_k - 1)(9n_k - 11)^2}$$

a_k = 1, -1

n_k = จำนวนคนในกลุ่มที่ k

J_k = จำนวนข้อในแบบทดสอบฉบับที่ k

$r_{\alpha k}$ = ค่าความเชื่อมั่นของแบบทดสอบฉบับที่ k

$$\sum_{k=1}^k a_k = 0$$

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและการแปลความหมายของผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ ดังนี้

n	แทน	จำนวนข้อในแบบทดสอบ
N	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
$\bar{X}$	แทน	ค่าเฉลี่ยของคะแนนแบบทดสอบ
$\bar{X} \%$	แทน	ค่าเฉลี่ยร้อยละของคะแนนแบบทดสอบ
S	แทน	ส่วนเบี่ยงเบนมาตรฐาน
CV	แทน	ค่าสัมประสิทธิ์การกระจาย
s_x^2	แทน	ความแปรปรวนของคะแนนทั้งหมด
$\sum S_i^2$	แทน	ผลรวมความแปรปรวนของคะแนนสอบแต่ละส่วน
$\sum x_i^2$	แทน	ผลรวมสัดส่วนของจำนวนข้อสอบแต่ละส่วนย่อย
M	แทน	การทดสอบด้วยสถิติ M. (M - test)

การเสนอผลการวิเคราะห์ข้อมูล

การศึกษาวิจัยครั้งนี้ ผู้วิจัยได้นำคะแนน ผลการสอบ มาวิเคราะห์ข้อมูลเป็นสี่ตอนเรียงตามลำดับดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ

ตอนที่ 2 ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบต้นฉบับที่คัดเลือกไว้แต่แรก จำนวน 40 ข้อ กับแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู

ตอนที่ 4 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ

การวิเคราะห์ข้อมูลครั้งนี้ ผู้วิจัยได้นำคะแนนผลการสอบของแบบทดสอบทั้ง 11 ฉบับของกลุ่มตัวอย่าง จำนวน 608 คน มาคำนวณค่าสถิติพื้นฐาน ซึ่งได้แก่ ค่าเฉลี่ย ค่าเฉลี่ยร้อยละ ส่วนเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย ของคะแนนแบบทดสอบในแต่ละฉบับ ปรากฏผลดังแสดงในตาราง 3

ตาราง 3 ค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ

แบบทดสอบ ฉบับที่	การลดจำนวนข้อ โดยใช้เกณฑ์	n	N	$\bar{X}$	S	$\bar{X}\%$	CV
1	-	40	67	20.9851	8.1928	52.4628	39.0410
2	ค่าสถิติ	35	56	15.2500	6.9955	43.5714	45.8721
3	ค่าสถิติ	30	54	12.2778	5.9547	40.9260	48.4997
4	ค่าสถิติ	25	60	12.5833	5.4433	50.3322	43.2581
5	ค่าสถิติ	20	51	8.2745	4.4471	41.3725	53.7446
6	ค่าสถิติ	15	50	4.6600	2.4673	31.0667	52.9464
7	ดุลพินิจของครู	35	64	17.4063	7.1174	49.7323	40.8898
8	ดุลพินิจของครู	30	55	14.3455	5.7847	47.8183	40.3241
9	ดุลพินิจของครู	25	51	13.3725	6.0914	53.4900	45.5517
10	ดุลพินิจของครู	20	50	9.3800	4.0768	46.9000	43.4627
11	ดุลพินิจของครู	15	50	7.7800	2.6111	51.8667	33.5617

ผลการวิเคราะห์ตามตาราง 3 ค่าสถิติพื้นฐานของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ พบว่าแบบทดสอบที่มีจำนวน 40, 35, 30, 25, 20 และ 15 ข้อที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีค่าเฉลี่ยตั้งแต่ 4.6600 ถึง 20.9851 ส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบมีค่าตั้งแต่ 2.4673 ถึง 8.1928 สำหรับแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู มีค่าเฉลี่ยตั้งแต่ 7.7800 ถึง

20.9851 ส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบมีค่าตั้งแต่ 2.6111 ถึง 8.1928 และเมื่อพิจารณาถึงค่าเฉลี่ยร้อยละกับค่าสัมประสิทธิ์การกระจายของแบบทดสอบแต่ละฉบับแล้ว ปรากฏว่า

แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา คะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ เป็นแบบทดสอบฉบับที่ง่ายที่สุด คือ มีค่าเฉลี่ยร้อยละ 50.3332 และแบบทดสอบที่มีจำนวน 15 ข้อเป็นแบบทดสอบฉบับที่ยากที่สุดคือ มีค่าเฉลี่ยร้อยละ 31.0667 ส่วนแบบทดสอบที่มีจำนวน 35, 30 และ 20 ข้อเป็นแบบทดสอบฉบับที่ค่อนข้างยากพอ ๆ กัน และแบบทดสอบทุกฉบับคือ แบบทดสอบที่มีจำนวน 35, 30, 25, 20 และ 15 ข้อ ยากกว่าแบบทดสอบฉบับที่มีจำนวน 40 ข้อ แบบทดสอบที่มีจำนวน 25 ข้อ ยากกว่าแบบทดสอบที่มีจำนวน 40 ข้อเล็กน้อย สำหรับค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า แบบทดสอบที่มีจำนวน 25 ข้อ มีการกระจายต่ำสุด คือ 43.2581 ซึ่งมีค่าใกล้เคียงกับแบบทดสอบที่มีจำนวน 40 ข้อมากที่สุด คือ 39.0410 และแบบทดสอบที่มีจำนวน 20 ข้อ มีการกระจายสูงที่สุด คือ 53.7446 โดยสรุปแล้วแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีคะแนนการกระจายส่วนใหญ่ใกล้เคียงกัน

แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ดุลพินิจของครูพิจารณาคะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ เป็นแบบทดสอบฉบับที่ง่ายที่สุด คือ มีค่าเฉลี่ยร้อยละ 53.4900 และแบบทดสอบที่มีจำนวน 20 ข้อเป็นแบบทดสอบฉบับที่ยากที่สุดคือ มีค่าเฉลี่ยร้อยละ 46.9000 นอกจากนั้น แบบทดสอบที่มีจำนวน 35, 30, และ 15 ข้อ เป็นแบบทดสอบฉบับที่ยากพอ ๆ กัน และแบบทดสอบที่มีจำนวน 25 ข้อและ 15 ข้อ เป็นแบบทดสอบฉบับที่ง่ายพอ ๆ กับแบบทดสอบฉบับที่มีจำนวน 40 ข้อ สำหรับค่าสัมประสิทธิ์การกระจาย จะเห็นว่า แบบทดสอบที่มีจำนวน 15 ข้อ มีการกระจายต่ำสุด คือ 33.5617 และแบบทดสอบที่มีจำนวน 25 ข้อมีการกระจายสูง ที่สุด คือ 45.5517 ส่วนแบบทดสอบที่มีจำนวน 30 ข้อ มีค่าสัมประสิทธิ์การกระจายที่ใกล้เคียงกับ แบบทดสอบที่มีจำนวน 40 ข้อ มากที่สุด โดยสรุปแล้วแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครูมีคะแนนการกระจายส่วนใหญ่ใกล้เคียงกัน

แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่จำนวนข้อเท่ากัน พบว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีค่าเฉลี่ยร้อยละต่ำกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู ทุกฉบับ และแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีการกระจายของคะแนนสูงกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู ยกเว้นแบบทดสอบที่มีจำนวน 25 ข้อ ที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีการกระจายของคะแนนต่ำกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู

ตอนที่ 2 ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบทั้ง 11 ฉบับ

ในการวิเคราะห์ข้อมูลครั้งนี้ เป็นการวิเคราะห์หาค่าความเชื่อมั่นของแบบทดสอบเลือก

ตอบทั้ง 11 ฉบับ ซึ่งจำแนกเป็นแบบทดสอบต้นฉบับ ที่คัดเลือกไว้แต่แรกจำนวน 40 ข้อ จากนั้น ได้ทำการลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู โดยการวิเคราะห์ครั้งนี้ ได้นำคะแนนผลการสอบจากกลุ่มตัวอย่างของแบบทดสอบแต่ละฉบับ มาแบ่งเป็น 3 ส่วน โดยใช้ อัตราส่วน 2:1:1 แล้วคำนวณหาค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้สูตรของเฟลด์ต์ (Feldt, 1975)

2.1 แบบทดสอบต้นฉบับ ที่คัดเลือกไว้แต่แรกจำนวน 40 ข้อ ปรากฏผลดังแสดงใน ตาราง 4

ตาราง 4 แสดงค่าความเชื่อมั่นของแบบทดสอบ ฉบับที่ 1 ที่มีจำนวน 40 ข้อ

แบบทดสอบ ฉบับที่	n	N	s_x^2	$\sum S_i^2$	$\sum x_i^2$	$r_{xx'}$
1	40	67	67.1224	28.8500	.3628	.8948

ผลการวิเคราะห์ตามตาราง 4 พบว่า แบบทดสอบฉบับที่ 1 ที่มีจำนวน 40 ข้อ มีผู้เข้าสอบ 67 คน มีค่าความแปรปรวน 67.1224 และมีค่าความเชื่อมั่นเท่ากับ .8948

2.2 แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา ปรากฏผลดังแสดงในตาราง 5

ตาราง 5 แสดงค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา

แบบทดสอบ ฉบับที่	n	N	s_x^2	$\sum S_i^2$	$\sum \mathcal{N}^2$	$r_{xx'}$
2	35	56	48.9374	28.4400	.3428	.6373
3	30	54	35.4579	16.0001	.3475	.8410
4	25	60	29.6292	14.1464	.3888	.8550
5	20	51	19.7765	8.9907	.3518	.8413
6	15	50	6.0874	4.0005	.3242	.5073

ผลการวิเคราะห์ตามตาราง 5 พบว่า แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา คือ ตัดข้อที่มีค่าความยากและค่าอำนาจจำแนกที่มีค่าน้อยที่สุดตามลำดับออกไปในแต่ละ จุดประสงค์ ได้แบบทดสอบจำนวน 5 ฉบับ ที่มีจำนวนข้อ 35, 30, 25, 20 และ 15 ข้อ มีค่าความแปรปรวนของคะแนนเท่ากับ 48.9374 , 35.4579 , 29.6292 , 19.7765 และ 6.0874 ตามลำดับ และ มีค่าความเชื่อมั่นเท่ากับ .6373 , .8410 , .8550 , .8413 และ .5073 ตามลำดับ จะเห็นว่าแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา คะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด และคะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 20 ข้อ มีค่าความเชื่อมั่นรองลงมา ส่วนคะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด

2.3 แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์คุณลักษณะของครูพิจารณา ปรากฏผลดังแสดงในตาราง 6

ตาราง 6 แสดงค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ดุลพินิจของครู
พิจารณา

แบบทดสอบ ฉบับที่	n	N	s_x^2	$\sum S_i^2$	$\sum A_i^2$	$r_{xx'}$
7	35	64	50.6574	24.9901	.3508	.7805
8	30	55	33.4628	15.5274	.3347	.8056
9	25	51	37.1051	16.0857	.3424	.8614
10	20	50	16.6204	8.2776	.3541	.7772
11	15	50	6.8179	4.7369	.3753	.4886

ผลการวิเคราะห์ตามตาราง 6 พบว่า แบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ดุลพินิจของครูพิจารณา คือ ข้อที่มีค่าดัชนีความสอดคล้องต่ำที่สุดออกไปในแต่ละจุดประสงค์ ได้แบบทดสอบจำนวน 5 ฉบับ ที่มีจำนวนข้อ 35, 30, 25, 20 และ 15 ข้อ มีค่าความแปรปรวนของคะแนนเท่ากับ 50.6574 , 33.4628 , 37.1015 , 16.6204 และ 6.8197 ตามลำดับ และมีค่าความเชื่อมั่นเท่ากับ .7805, .8056 , .8614 , .7772 และ .4886 ตามลำดับ จะเห็นว่า แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ดุลพินิจของครูพิจารณา คะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด และคะแนนที่ได้จากการทดสอบ ด้วยแบบทดสอบที่มีจำนวน 30 ข้อ มีค่าความเชื่อมั่นรองลงมา ส่วนแบบทดสอบที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด

ตอนที่ 3 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบต้นฉบับ ที่คัดเลือกไว้แต่แรก
จำนวน 40 ข้อ กับแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู

3.1 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับพร้อมกัน

การวิเคราะห์ข้อมูลครั้งนี้ ได้นำค่าความเชื่อมั่นของแบบทดสอบทั้ง 11 ฉบับ ซึ่งเป็นแบบทดสอบต้นฉบับ ที่คัดเลือกไว้แต่แรกจำนวน 40 ข้อ 1 ฉบับ แบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ค่าสถิติ 5 ฉบับ และโดยใช้เกณฑ์ดุลพินิจของครูอีก 5 ฉบับ มาทดสอบความแตกต่างกัน โดยอาศัยสูตรของ มาราสคูโย (Marascuilo) ปรากฏผลดังแสดงในตาราง 7

ตาราง 7 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบทั้ง 11 ฉบับ

แบบทดสอบ ฉบับที่	การลดจำนวนข้อ โดยใช้เกณฑ์	n	N	$r_{xx'}$	M
1	-	40	67	.8948	77.5951**
2	ค่าสถิติ	35	56	.6373	
3	ค่าสถิติ	30	54	.8410	
4	ค่าสถิติ	25	60	.8550	
5	ค่าสถิติ	20	51	.8413	
6	ค่าสถิติ	15	50	.5073	
7	ดุลพินิจของครู	35	64	.7805	
8	ดุลพินิจของครู	30	55	.8056	
9	ดุลพินิจของครู	25	51	.8614	
10	ดุลพินิจของครู	20	50	.7772	
11	ดุลพินิจของครู	15	50	.4886	

$$\chi^2(\alpha = .01, df = 10) = 23.209$$

ผลการวิเคราะห์ตามตาราง 7 จากการเปรียบเทียบค่าความเชื่อมั่นแต่ละคู่โดยใช้สูตรมาราสกุโย (Marascuilo) พบว่า ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์สถิติ และโดยใช้เกณฑ์ดุลพินิจของครู แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าค่าความเชื่อมั่นของแบบทดสอบทั้ง 11 ฉบับ ต้องมีอย่างน้อย 1 คู่ มีความแตกต่างกัน

3.2 การเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบต้นฉบับ ที่คัดเลือกไว้แต่แรก จำนวน 40 ข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู เป็นรายคู่ โดยอาศัยสูตรของเชฟเฟ้ (Scheffe') ปรากฏผลดังแสดงในตาราง 8

ตาราง 8 การเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงทั้ง 2 วิธี เป็นรายคู่

แบบทดสอบฉบับที่	จำนวนข้อ	ความแตกต่างเป็นรายคู่ ($\sqrt{F}$)
1 กับ 2	40 กับ 35	20.0086 **
1 กับ 3	40 กับ 30	2.3160
1 กับ 4	40 กับ 25	.9164
1 กับ 5	40 กับ 20	2.1833
1 กับ 6	40 กับ 15	25.8803 **
1 กับ 7	40 กับ 35	8.1844 **
1 กับ 8	40 กับ 30	5.2236 *
1 กับ 9	40 กับ 25	.9815
1 กับ 10	40 กับ 20	6.9165 **
1 กับ 11	40 กับ 15	27.5184 **

$$\sqrt{F}(\alpha = .05, df = 1) = 3.841^*, \quad \sqrt{F}(\alpha = .01, df = 1) = 6.635^{**}$$

ผลการวิเคราะห์ตามตาราง 8 พบว่า การเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ (ฉบับที่ 1) กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ค่าสถิติ ได้ค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 2 คู่ ได้แก่ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับ 35 ข้อ (ฉบับที่ 1 กับ 2) และจำนวน 40 ข้อ กับ 15 ข้อ (ฉบับที่ 1 กับ 6) และได้ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ที่ระดับ .05 จำนวน 3 คู่ ได้แก่ ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับ 30 ข้อ (ฉบับที่ 1 กับ 3) จำนวน 40 ข้อ กับ 25 ข้อ (ฉบับที่ 1 กับ 4) และจำนวน 40 ข้อ กับ 20 ข้อ (ฉบับที่ 1 กับ 5) และจากการเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ (ฉบับที่ 1) กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์จุดพินิจของครู ได้ค่าความเชื่อมั่นของแบบทดสอบแตกต่าง

กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 3 คู่ ได้แก่ ค่าความเชื่อมั่นของแบบทดสอบ ที่มีจำนวน 40 ข้อ กับ 35 ข้อ (ฉบับที่ 1 กับ 7) จำนวน 40 ข้อ กับ 20 ข้อ (ฉบับที่ 1 กับ 10) และจำนวน 40 ข้อ กับ 15 ข้อ (ฉบับที่ 1 กับ 11) ได้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับ 30 ข้อ (ฉบับที่ 1 กับ 8) และได้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 1 คู่ ได้แก่ ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับ 25 ข้อ โดยสรุปแล้วแบบทดสอบที่มีจำนวน 25 ข้อ ที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครูมีค่า $\sqrt{F}$ ต่ำ น่าจะใช้แทนแบบทดสอบที่มีจำนวน 40 ข้อดีที่สุด

ข้อที่ 4 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน

การวิเคราะห์ครั้งนี้ ได้นำค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน มาทดสอบความแตกต่างเป็นรายคู่ ด้วยสูตรของ เชฟเฟ (Scheffe') ปรากฏผลดังแสดงในตาราง 9

ตาราง 9 การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน

แบบทดสอบฉบับที่	จำนวนข้อ	ความแตกต่างรายคู่ ($\sqrt{F}$)
2 กับ 7	35	3.5588
3 กับ 8	30	.5208
4 กับ 9	25	.0173
5 กับ 10	20	1.7485
6 กับ 11	15	.0158

$$\sqrt{F}(\alpha = .05, df = 1) = 3.8$$

ผลการวิเคราะห์ตามตาราง 9 พบว่า จากการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่า

กันเป็นรายคู่ ได้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ที่ระดับ .05 ทุกคู่ แสดงว่า แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ให้ค่าความเชื่อมั่นไม่แตกต่างกันทุกขนาดความยาว

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

การวิจัยครั้งนี้ เป็นการศึกษาการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู โดยมีจุดมุ่งหมายดังต่อไปนี้

1. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ
2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับ ฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครู
3. เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดสุพรรณบุรี จำนวน 608 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ด้านโจทย์ปัญหา ชั้นประถมศึกษาปีที่ 5 จำนวน 11 ฉบับ ซึ่งผู้วิจัยได้สร้างขึ้นให้ครอบคลุมเนื้อหาเรื่องโจทย์ปัญหาเศษส่วน โจทย์ปัญหาคณนิยม และโจทย์ปัญหาร้อยละ โดยมีจำนวนข้อสอบในแบบทดสอบแต่ละฉบับที่ลดลงจากฉบับ 40 ข้อ เป็น ฉบับ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครูในการคัดเลือกข้อสอบออก

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการตามขั้นตอน ดังนี้

1. นำแบบทดสอบไปดำเนินการทดสอบตามโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างที่ได้ขออนุญาตไว้

2. ดำเนินการสอบ โดยอธิบายให้ผู้สอบทุกคนเข้าใจถึงวัตถุประสงค์และผลประโยชน์ที่ได้รับจากการตอบแบบทดสอบ
3. นำกระดาษคำตอบมาตรวจให้คะแนน
4. นำคะแนนมาวิเคราะห์ทางสถิติและทดสอบตามสมมติฐานที่ตั้งไว้

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้วิเคราะห์ข้อมูลตามลำดับขั้น ดังนี้

1. หาค่าสถิติพื้นฐานของแบบทดสอบทั้ง 11 ฉบับ
2. หาค่าความเชื่อมั่นของแบบทดสอบทั้ง 11 ฉบับ
3. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับ ฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู
4. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน

สรุปผลการวิจัย

ผลการวิจัยสรุปได้ดังนี้

1. แบบทดสอบ

1.1 **แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติพิจารณา** เมื่อพิจารณาถึงค่าเฉลี่ยร้อยละของแบบทดสอบแต่ละฉบับแล้ว แบบทดสอบที่มีจำนวน 25 ข้อ เป็นแบบทดสอบฉบับที่ง่ายที่สุด และแบบทดสอบที่มีจำนวน 15 ข้อ เป็นแบบทดสอบฉบับที่ยากที่สุด ส่วนแบบทดสอบที่มีจำนวน 35, 30 และ 20 ข้อเป็นแบบทดสอบฉบับที่ค่อนข้างยากพอ ๆ กันและแบบทดสอบทุกฉบับ คือ แบบทดสอบที่มีจำนวน 35, 30, 25, 20 และ 15 ข้อ ยากกว่าแบบทดสอบที่มีจำนวน 40 ข้อ แบบทดสอบที่มีจำนวน 25 ข้อ ยากกว่าแบบทดสอบที่มีจำนวน 40 ข้อ เล็กน้อย และเมื่อพิจารณาถึง ค่าสัมประสิทธิ์การกระจายของแบบทดสอบแต่ละฉบับแล้ว แบบทดสอบที่มีจำนวน 25 ข้อ มีการกระจายต่ำสุด ซึ่งมีค่าใกล้เคียงกับแบบทดสอบที่มีจำนวน 40 ข้อ มากที่สุด และแบบทดสอบที่มีจำนวน 20 ข้อ มีการกระจายสูงที่สุด โดยสรุปแล้ว แบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติ ส่วนใหญ่ มีคะแนนการกระจายใกล้เคียงกัน

1.2 **แบบทดสอบที่ลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครูพิจารณา** เมื่อพิจารณาถึงค่าเฉลี่ยร้อยละของแบบทดสอบแต่ละฉบับแล้ว แบบทดสอบที่มีจำนวน 25 ข้อเป็น

แบบทดสอบฉบับที่ง่ายที่สุดและแบบทดสอบที่มีจำนวน 20 ข้อ เป็นแบบทดสอบฉบับที่ยากที่สุด นอกนั้น แบบทดสอบที่มีจำนวน 35, 30 และ 15 ข้อ เป็นแบบทดสอบฉบับที่ยากง่ายพอ ๆ กัน และแบบทดสอบที่มีจำนวน 25 ข้อ และ 15 ข้อ เป็นแบบทดสอบฉบับที่ง่ายพอ ๆ กับแบบทดสอบฉบับที่มีจำนวน 40 ข้อ และเมื่อพิจารณาถึงค่าสัมประสิทธิ์การกระจายของแบบทดสอบแต่ละฉบับแล้ว แบบทดสอบ ที่มีจำนวน 15 ข้อ มีการกระจายต่ำสุด และแบบทดสอบที่มีจำนวน 25 ข้อ มีการกระจายสูงที่สุด ส่วนแบบทดสอบที่มีจำนวน 30 ข้อ มีค่าสัมประสิทธิ์การกระจายที่ใกล้เคียงกับแบบทดสอบที่มีจำนวน 40 ข้อ มากที่สุด โดยสรุปแล้วแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครูส่วนใหญ่ มีคะแนนการกระจายใกล้เคียงกัน

1.3 แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน คือแบบทดสอบ ที่มีจำนวน 35, 30, 25, 20 และ 15 ข้อ ที่มีจำนวนข้อเท่ากัน แบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติ มีค่าเฉลี่ยร้อยละยากกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู และแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติมีการกระจายของคะแนนสูงกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู ยกเว้นแบบทดสอบที่มีจำนวน 25 ข้อ ที่คัดเลือกโดยใช้เกณฑ์ค่าสถิติ มีการกระจายของคะแนนต่ำกว่าแบบทดสอบที่คัดเลือกโดยใช้เกณฑ์ดุลพินิจของครู

2. ค่าความเชื่อมั่นของแบบทดสอบ

จากการนำคะแนนผลการสอบจากกลุ่มตัวอย่าง ของแบบทดสอบแต่ละฉบับมาแบ่งเป็น 3 ส่วน โดยใช้อัตราส่วน 2:1:1 แล้วคำนวณหาค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้สูตรของ เฟลด์ต์ (Feldt, 1975) (บุญเว็ด วิทยุอนันตพงษ์, 2533:) ได้ผลดังนี้

2.1 ค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่มี จำนวนข้อลดลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ พิจารณา คะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนคะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด และจากการเปรียบเทียบ ค่าความเชื่อมั่นแต่ละฉบับ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จึงได้นำไปทดสอบความแตกต่างรายคู่ พบว่าแตกต่างเฉพาะค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับ แบบทดสอบที่ลดจำนวนข้อลงเหลือ 35 และ 15 ข้อ ที่เหลืออื่นนั้นไม่พบความแตกต่าง

2.2 ค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับ ที่ลดจำนวนข้อลงเหลือ 35 , 30 , 25 , 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครู

พิจารณา คะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนคะแนนที่ได้จากการทดสอบด้วยแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด และจากการเปรียบเทียบ ค่าความเชื่อมั่นแต่ละฉบับ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ต่อจากนั้น จึงได้ นำไปทดสอบความแตกต่างรายคู่ พบว่าแตกต่างกันเป็นส่วนใหญ่ ยกเว้นค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 25 ข้อ เพียงฉบับเดียวที่ให้ค่าความเชื่อมั่นไม่แตกต่างกัน

2.3 ค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน เมื่อนำแบบทดสอบมาทดสอบความแตกต่างรายคู่ พบว่าได้ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่ ดังนั้น แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ให้ค่าความเชื่อมั่น ไม่แตกต่างกัน

อภิปรายผล

การวิจัยครั้งนี้ มีเป้าหมายสำคัญ เพื่อค้นหาว่าแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อลงโดยใช้เกณฑ์สถิติ กับใช้เกณฑ์ดุลพินิจของครู ว่าความยาวของแบบทดสอบ เมื่อลดจำนวนข้อให้สั้นลง ให้มีจำนวนน้อยข้อนั้น ควรจะมีจำนวนข้อเท่าใดจึงจะทำให้ค่าความเชื่อมั่นเท่าเดิมซึ่งสามารถอภิปรายตามสมมติฐาน ได้ดังนี้

1. แบบทดสอบที่ใช้เกณฑ์ค่าสถิติ ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ มีค่าความเชื่อมั่นแตกต่างจากแบบทดสอบฉบับที่มีจำนวน 40 ข้อ

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบฉบับที่มีจำนวน 40 ข้อ กับ ค่าความเชื่อมั่นของฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติพิจารณา พบว่าแบบทดสอบที่มีจำนวน 40 ข้อ มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 35 และ 15 ข้อซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 1 ทั้งนี้เมื่อพิจารณาจากผลการศึกษา จะพบว่าแบบทดสอบที่ลดจำนวนข้อลงเหลือ 35 ข้อ และ 15 ข้อ มีค่าความเชื่อมั่นต่ำ อาจเนื่องมาจากความบกพร่องในการดำเนินการสอบ ผู้วิจัยได้สุมนักเรียนแต่ละห้องออกเป็น 2 – 4 กลุ่ม แต่ละกลุ่มจะได้แบบทดสอบที่มีจำนวนข้อไม่เท่ากัน ซึ่งนักเรียนแต่ละกลุ่มต่าง ๆ ต้องสอบในห้องเดียวกัน นักเรียน ที่ตอบข้อสอบที่มีจำนวนน้อยกว่าจะตอบเสร็จก่อน นักเรียนที่ตอบข้อสอบซึ่งมีจำนวนข้อมากกว่า เมื่อเห็นเพื่อน ๆ ตอบเสร็จก่อน ก็มีความตั้งใจในการสอบลดลง อาจจะรับ

ตอบให้เสร็จโดยเร็ว ทำให้เกิดความคลาดเคลื่อนในการวัดขึ้น (ยศอารีย์ รวยธนพานิช, 2534 : 58) มีผลทำให้ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 35 ข้อ ต่ำกว่าค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40, 30, 25 และ 20 ข้อ ซึ่งสอดคล้องกับการศึกษาของ ผดุงชาติ สุวรรณวงศ์. (2516 : 42 – 43) ที่พบว่า มีข้อสอบบางฉบับเมื่อเพิ่มความยาวขึ้นแล้วมีค่าความเชื่อมั่นต่ำลง ส่วนค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ มีค่าแตกต่างจากค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 15 ข้อ อาจเป็นเพราะจำนวนของแบบทดสอบที่สั้นเกินไป จะมีผลทำให้ค่าความเชื่อมั่นมีค่าต่ำกว่าแบบทดสอบที่มีจำนวนข้อมากกว่า ซึ่งสอดคล้องกับที่ กรอนลันด์ (Gronlund, 1976 : 117) และ เมห์เรนส์ และเลห์แมนน์ (Mehrens and Lehmann, 1984 : 278 – 279) กล่าวไว้ว่า แบบทดสอบ ที่มีจำนวนข้อมากจะมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มีจำนวนข้อน้อย ส่วนแบบทดสอบ ที่มีจำนวน 40 ข้อ มีความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ที่ระดับ .05 กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 30, 25 และ 20 ข้อ ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อที่ 1 ทั้งนี้เมื่อพิจารณาผลการศึกษาค้นพบว่าแบบทดสอบที่ลดจำนวนข้อลงเหลือ 30, 25 และ 20 ข้อ มีค่าความเชื่อมั่นสูงใกล้เคียงกับค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ และค่าความเชื่อมั่นที่สูงใกล้เคียงที่สุด คือ แบบทดสอบที่มีจำนวน 25 ข้อ ค่าความเชื่อมั่นที่ต่ำที่สุด คือ แบบทดสอบที่มีจำนวน 15 ข้อ อาจเป็นเพราะแบบทดสอบที่มีจำนวน 25 ข้อ เป็นแบบทดสอบฉบับที่ยากกว่าแบบทดสอบที่มีจำนวน 40 ข้อเพียงเล็กน้อย และค่าสัมประสิทธิ์การกระจาย ก็มีค่าใกล้เคียงมากที่สุด กับแบบทดสอบที่มีจำนวน 40 ข้อ จึงทำให้ค่าความเชื่อมั่นสูงใกล้เคียงที่สุดกับแบบทดสอบที่มีจำนวน 40 ข้อ ส่วนแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำที่สุดนั้น อาจเป็นเพราะ แบบทดสอบที่มีจำนวน 15 ข้อ เป็นแบบทดสอบฉบับที่ยากที่สุด และค่าความแปรปรวนของคะแนนของแบบทดสอบมีค่าต่ำสุด จึงทำให้ค่าความเชื่อมั่น มีค่าต่ำสุด ซึ่งสอดคล้องจากคำกล่าวของเมห์เรนส์ และ เลห์แมนน์ (Mehrens and Lehmann, 1984 : 278 – 281) ที่กล่าวถึงองค์ประกอบข้อหนึ่งที่มีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบ คือ ความแปรปรวนระหว่างคะแนนผลการสอบ ถ้าความแปรปรวนของคะแนนมีน้อย ค่าความเชื่อมั่นก็ลดลง แต่ถ้าความแปรปรวนของคะแนนมีมาก ก็จะทำให้ค่าความเชื่อมั่นสูงขึ้น

2. แบบทดสอบที่ใช้เกณฑ์ดุลพินิจของครู ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ มีค่าความเชื่อมั่นแตกต่างจากแบบทดสอบฉบับที่มีจำนวน 40 ข้อ

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบฉบับที่มีจำนวน 40 ข้อ กับค่าความเชื่อมั่นของฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจ

ของครู พิจารณา พบว่า แบบทดสอบที่มีจำนวน 40 ข้อ มีค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติ กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง เหลือ 35, 30, 20 และ 15 ข้อ ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 ทั้งนี้เนื่องจากแบบทดสอบที่มีความยาวมาก ต้องใช้เวลานาน ทำให้ต้องอ่านข้อสอบเป็นจำนวนมาก ทำให้เกิดความเบื่อหน่ายจึงเป็นเหตุให้ผู้ตอบไม่ตั้งใจตอบเป็นจำนวนมาก ตรงกันข้ามกับเมื่อแบบทดสอบมีความยาวลดลง จะมีผู้ตั้งใจตอบแบบทดสอบมากขึ้น (อรนุช อธิรักษพานิชย์, 2538 : 92) ซึ่งสอดคล้องกับที่บุญเรือน จรัสวิมล. (2514 : 14) กล่าวไว้ว่า เมื่อเพิ่มจำนวนข้อสอบให้มากขึ้น นักเรียนต้องใช้เวลานานขึ้น มักจะเกิดความเหนื่อยล้าและ เบื่อหน่ายในการทำข้อสอบ ส่วนแบบทดสอบที่มีจำนวน 40 ข้อ มีค่าความเชื่อมั่นแตกต่างกัน อย่าง ไม่มีนัยสำคัญทางสถิติ ที่ระดับ .05 กับค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลงเหลือ 25 ข้อ ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 ทั้งนี้เมื่อพิจารณาผลการศึกษา จะพบว่า แบบทดสอบที่ลดจำนวนข้อลงเหลือ 25 ข้อ มีค่าความเชื่อมั่นสูงใกล้เคียงกับค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 40 ข้อ และแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด ทั้งนี้ อาจเป็นเพราะค่าสัมประสิทธิ์การกระจายของแบบทดสอบที่มีจำนวน 25 ข้อ มีการกระจายสูงที่สุด แบบทดสอบ ที่มีจำนวน 15 ข้อ มีการกระจายต่ำสุด และค่าความแปรปรวนของคะแนนของแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าต่ำสุด จึงทำให้ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 25 ข้อมีค่าสูงสุด และ ค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าต่ำสุด ซึ่งสอดคล้องจากคำกล่าวของกรอนลันด์ (Gronlund, 1976 : 119) ที่ว่าค่าความเชื่อมั่นย่อมขึ้นอยู่กับค่าความแปรปรวนระหว่างคะแนน ผลการสอบ ถ้าความแปรปรวนระหว่างคะแนนผลการสอบมีน้อย ย่อมได้ค่าความเชื่อมั่นต่ำกว่า คะแนนผลการสอบที่มีความแปรปรวนมาก

3. แบบทดสอบที่ลดจำนวนข้อลงโดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครูที่มีจำนวนข้อเท่ากัน มีค่าความเชื่อมั่นแตกต่างกัน

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน พบว่า แบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติระดับ .05 ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อที่ 3 ทั้งนี้เมื่อพิจารณาคะแนนผลการสอบจากกลุ่มตัวอย่างของแบบทดสอบในแต่ละฉบับ ทั้งโดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากันแล้ว ส่วนใหญ่จะมีค่าความแปรปรวนใกล้เคียงกัน จึงทำให้ค่าความเชื่อมั่นไม่แตกต่างกัน ซึ่งสอดคล้องกับคำกล่าวของ กรอนลันด์

(Gronlund. 1976 : 119) ที่ว่าค่าความเชื่อมั่นย่อมขึ้นอยู่กับค่าความแปรปรวนระหว่างคะแนนผลการสอบ ถ้าความแปรปรวนของคะแนนผลการสอบมีน้อยย่อมได้ค่าความเชื่อมั่นต่ำกว่าคะแนนผลการสอบที่มีความแปรปรวนมากดังนั้น แบบทดสอบที่มีความแปรปรวนใกล้เคียงกัน ย่อมมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ข้อเสนอแนะ

ผลจากการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้ และในการทำวิจัย ต่อไปดังนี้

1. ข้อเสนอแนะในการนำไปใช้

1.1 ในการนำแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ด้านโจทย์ปัญหาไปใช้กับนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดสุพรรณบุรี เรื่องทศนิยม เศษส่วน และโจทย์ปัญหาร้อยละ ควรมีจำนวน 25 ข้อก็เพียงพอแทนที่จะมีถึงจำนวน 40 ข้อ เพราะจะช่วยประหยัดเวลาและแรงงานทั้งผู้สอบและผู้ตรวจ โดยที่มีค่าความเชื่อมั่นเหมือนเดิม

1.2 แบบทดสอบเลือกตอบ เมื่อจะลดจำนวนข้อให้น้อยลง จะคัดเลือกข้อสอบออกโดยใช้เกณฑ์ค่าสถิติ หรือโดยใช้เกณฑ์ดุลพินิจของครูก็ได้ แต่ถ้าต้องการประหยัดค่าใช้จ่ายและเวลาสามารถใช้เกณฑ์ดุลพินิจของครูโดยใช้ครูที่มีประสบการณ์สูงในวิชานั้นประมาณ 5 คน พิจารณาร่วมกัน

2. ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาจำนวนข้อที่เหมาะสมของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนในวิชาต่าง ๆ และกับระดับชั้นอื่น ๆ เพื่อเปรียบเทียบกับผลการศึกษาค้างนี้และ เพื่อประโยชน์ในการสร้างแบบทดสอบเลือกตอบที่มีคุณลักษณะที่ดีในแง่ประสิทธิภาพและความเพียงพอของข้อคำถาม

2.2 ควรมีการศึกษาค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่มีความยาวขนาดต่าง ๆ กัน ที่มีหลายส่วนโดยใช้สูตรของเฟลลด์

2.3 ควรมีการศึกษาถึงการวิเคราะห์ตัวลงที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ

บรรณานุกรม

บรรณานุกรม

- กาญจนา มณีแสง. การประเมินผลทางการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2523.
- ชนะ ทานะวงศ์. การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ชนิดเลือกตอบที่มีจำนวนข้อและตัวเลือกแตกต่างกัน วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 4. ปรินญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. อัดสำเนา.
- ชวาล แพรัตกุล. เทคนิคการเขียนคำถามเลือกตอบ. กรุงเทพฯ : กิจจันทร์การพิมพ์, 2530. _____ เทคนิคการวัดผล. พิมพ์ครั้งที่ 5. กรุงเทพฯ : วัฒนาพานิช, 2516.
- ดร.น. ชาญตระกูล. การเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 7 ที่มีจำนวนตัวเลือกไม่เท่ากัน. ปรินญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519. อัดสำเนา.
- เดือนใจ เกตุษา. การสร้างแบบทดสอบ 1. พิมพ์ครั้งที่ 3. กรุงเทพฯ : หจก. คุณพิน อักษรกิจ, 2529.
- บุญเชิด ภิญโญอนันตพงษ์. การวัดและการประเมินผลการศึกษา : ทฤษฎีและการประยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพฯ : ภาควิชาพื้นฐานการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. _____ การสร้างแบบทดสอบ 1. ภาควิชาพื้นฐานการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2525. _____ การทดสอบแบบอิงเกณฑ์ : แนวคิด และปฏิบัติ. คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2527. _____ "Congeneric part Reliability," วารสารการวัดผล การศึกษา. 12(34) ; พฤษภาคม - สิงหาคม 2533. _____ "การวิเคราะห์ข้อสอบอิงจุดประสงค์ชนิดครูสร้าง; วิธีการใหม่," วารสารการวัดผลทางการศึกษา. 13(39): 12-32; มกราคม-เมษายน 2535.

- บุญเจิด ภิญโญอนันตพงษ์. "เทคนิคการประมาณค่าความเชื่อถือได้ของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์," วารสารการวัดผลการศึกษา 14(42) : 8-20 ; มกราคม-เมษายน 2536.
- บุญเรียง ขจรศิลป์. วิธีวิจัยทางการศึกษา. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2533.
- บุญเรือน จรัสวิมล. การเพิ่มข้อสอบที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ. ปรินญาณิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. อัดสำเนา.
- ผดุงชาติ สุวรรณวงศ์. "Reliability and validity ของข้อสอบที่เขียนเพิ่มให้ยาวขึ้น," ใน พัฒนาวัดผล 9. หน้า 42-43. กรุงเทพฯ : กรุงเทพมหานคร, 2516.
- พวงรัตน์ ทวีรัตน์. การสร้างและพัฒนาแบบทดสอบผลสัมฤทธิ์. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2529.
- ไพศาล หวังพานิช. การวัดผลการศึกษา กรุงเทพฯ : ไทยวัฒนาพานิช, 2526.
- ยศอารีย์ รวยธนพานิช. การเปรียบเทียบความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบที่มีช่วงความยากและความยาวของแบบทดสอบต่างกันในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3. ปรินญาณิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2534. อัดสำเนา.
- ล้วน สายยศ. "การวัดผลสัมฤทธิ์วิชาคณิตศาสตร์," พัฒนาวัดผล 11. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2518.
- "ความเชื่อมั่นอีกวิธีหนึ่ง," พัฒนาวัดผล 12. หน้า 76-82 กรุงเทพฯ : โคมการพิมพ์, 2519 .
- วารุณี ปิครวัชชัย. การศึกษาเปรียบเทียบค่าสถิติของแบบทดสอบชนิดเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน. ปรินญาณิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2513. อัดสำเนา.
- วิเชียร เกตุสิงห์. "วิธีวิเคราะห์องค์ประกอบกับการเลือกข้อสอบ," วารสารวิจัยการศึกษา 3(1) : มกราคม-กุมภาพันธ์ 2523.
- ศึกษาธิการ, กระทรวง. หลักสูตรประถมศึกษา พุทธศักราช 2521. ฉบับปรับปรุง พ.ศ. 2533. พิมพ์ครั้งที่ 2. กรุงเทพฯ : โรงพิมพ์การศาสนา. 2535.
- คู่มือครูคณิตศาสตร์ชั้นประถมศึกษาปีที่ 5. โรงพิมพ์คุรุสภา, พิมพ์ครั้งที่ 2. 2535.

- ศึกษาธิการ, กระทรวง. เครื่องมือวัดผลการเรียนรู้ตามจุดประสงค์การเรียนรู้. ในสมุด
ประจำชั้น ประถมศึกษาปีที่ 5. กรุงเทพฯ : โรงพิมพ์คุรุสภา, 2525.
- เสริมศักดิ์ วิชาลาภรณ์ และเอนกกุล กริแสง. หลักเบื้องต้นของการวัดผลการศึกษา. พิมพ์ครั้งที่ 3.
กรุงเทพฯ : โรงพิมพ์อักษรสัมพันธ์, 2519.
- สำเร็จ บุญเรืองรัตน์. ทฤษฎีการวัดและประเมินผลการศึกษา. พิมพ์ครั้งที่ 2 กรุงเทพฯ :
สยามศึกษา, 2529.
- อนันต์ ศรีโสภณ. ทฤษฎีการวัดและการทดสอบ. กรุงเทพฯ : ไทยวัฒนาพานิช, 2525.
- อรนุช อิตติรักษพานิชย์. การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ วัดบุคคลิกภาพ
ของเยาวชนไทย เมื่อลดความยาวให้สั้นลง. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ :
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2538. อัดสำเนา.
- Adams, Georgia S. Measurement and Evaluation in Education Psychology and Guidance. New
York : Holt Rinehart and Winstone, Inc., 1964.
- Aiken, Lewis R. "Testing with Multiple-choice Items," Journal of Research and Development in
Education. 20:44-58, Summer, 1987.
- Allen, Mary J. and Wendy M. Yen. Introduction to Measurement Theory. California .
Brooks/Cole, 1979.
- Anastasi, Ann. Psychological Testing. New York : Mc. Graw-Hill, Inc., 1968.
- Bloom, Benjamis s. and other s. Handbook of Formative and Sommative Evaluation of Student
Learning. New York : Mc Graw-Hill Book Compang, 1971.
- Budescu, David V. and Barch Nevo. "Optional Number of Options : An Investigation of the
Assumption of Proportionalify," Journal of Educational Measurement. 22:183-196;
Fall, 1985.
- Campbell, Alison C. "Some Determinant of Difficuty of Non-Verbal Classification Item,"
Educational and Psychological Measurement, 21 (Winter 1961) : 899 - 913.
- Ebel, Robert L. Essentials of Educational Measurement. 3rd ed. Englewood Cliffs : Prentice-
Hall, 1979.
- Fan, Chung-Teh. "Item Analysis Table," พิมพ์ในประเทศไทยจากต้นฉบับของ Edvcational
Testing Service, U.S.A. 1952, 32PF.

- Ferguson, George A. Statistical Analysis in Psychology and Education. 3rd ed. New York : Mc Graw-Hill Book Company, 1966.
- Forsyth, Robert A. and Ansley N. Timothy. "The Importance of computational Skill For Answering Items in a Mathematics Problem Solving test : Implications for construct Validity," Educational and Psychological Measurement. 42:257-263; 1982.
- Forsyth, Robert A. and Spratt F. Kouin. "Measuring Problem-Solving Ability in Mathematical With Multiple Choice Items : The Effect of Item Formation Selected Item and Test Characteristics," Journal of Educational Measurement. 17:31-43; 1980.
- Gronlund, Norman E. Measurement and Evaluation in Teaching. 3rd ed. New York : Macmillan, 1976.
- Hakstian, A. Ralph. and Whalen, Thomas E. "A K-Sample Significance Test for Independent Alpha Coefficients," Psychometrika. 41 : 219-231; 1976.
- Henryssen, S. "Gathering, Analyzing and Using Data on Test Items," Educational Measurement, edited by Thorndike, R.L. p. 130-159. Washington, D.C. : American Council on Educational, 1971.
- Hopkins, Charles D. and Antes, Richard L. Classroom Testing Construction. Illinois, F.E. Peacock Publishers, Inc., 1979.
- Krulik Stephen "Problem. Problem Solving and Strategy Games," The Mathematic Teacher. 8 : 649 - 651; November, 1977.
- Le Blane, F. "You Can Teach Problem Solving," The Arithmetic Teacher. 25 : 16 - 20; November, 1977.
- Lindvall, C. M. and Anthony J. Nitko. Measuring Pupil Achievement and Aptitude. New York : Harcourt Brace Jovanovich, Inc., 1967.
- Lord, Frederic M. and Melvin R. Novick. Statistical Theories of Mental Test Scores. California : Harcourt Brace Jovanovich, Inc., 1967.
- Magnusson, D. Test Theory. California : Addison-Wesley Publishing Company, 1967.
- Mehren, W. A. and Irvin J. Lehmann. Measurement and Evaluation in Education and Psychology. 3rd ed. Japan: CBS College Publishing, 1984.

- Nunnally, Jum C. Educational Measurement and Evaluation. New York : McGraw-Hill Book Company, 1964.
- Rimland, Bernard. "The effect of Varying Time Limits and of Using Right Answer not Given in Experimental of the U.S. Navy Arithmetic Test," Educational and Psychological Measurement. 20 : 533 - 538, Autumm, 1960.
- Stanley, Julian C. and Kenneth D. Hopkins. Educational and Psychology Evaluation. 5th ed. Engliwoods Cliff : Pretice-Hall, 1972.
- Tollefson, Nana, "A Comparison of The Item Difficulty and Item Discrimination of Multiple-choice Items Using the "none of the above" and one correct," Educational and Psychological Measurement. 47, 377 - 383, Summer, 1987.
- West, Tommic A. "Rx for Verbal Problem : A Diagrastic-Precriptive Approcach," Arithmetic Teacher. 25 : 57 - 58; 1977.
- Whirl Robert T. "Problem Solving-Solution or Technique," The Arithematies Teacher. 6 : 1551 - A : October, 1973.
- Zalewski Claire Jean. "An Investigation of Selected Factors, Contibuting to Success in Solving Mathematical Word Problem," Dissertation Abstracts International. 4 : 2804 - A; July 1978.

ภาคผนวก

ภาคผนวก ก

1. ค่าดัชนีความสอดคล้อง (IOC) ระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม
2. ค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติ

ตาราง 10 แสดงค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม (IOC)

ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC
1	1.00	21	1.00	41	1.00
2	1.00	22	1.00	42	1.00
3	1.00	23	1.00	43	1.00
4	1.00	24	1.00	44	1.00
5	1.00	25	1.00	45	1.00
6	1.00	26	1.00	46	1.00
7	1.00	27	1.00	47	1.00
8	1.00	28	1.00	48	1.00
9	1.00	29	1.00	49	1.00
10	1.00	30	1.00	50	1.00
11	1.00	31	1.00	51	1.00
12	1.00	32	1.00	52	1.00
13	1.00	33	1.00	53	1.00
14	1.00	34	1.00	54	1.00
15	1.00	35	1.00	55	1.00
16	1.00	36	0.60	56	1.00
17	1.00	37	1.00	57	1.00
18	1.00	38	1.00	58	1.00
19	1.00	39	1.00	59	1.00
20	1.00	40	1.00	60	1.00

ตาราง 11 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเป็นรายข้อ
ที่คัดเลือกไว้ จำนวน 40 ข้อ (ฉบับที่ 1)

ข้อที่	P _H	P _L	P	r	ข้อที่	P _H	P _L	P	r
1	1.00	.95	.74	.70	21	1.00	.68	.84	.47
2	1.00	0.00	.50	1.00	22	1.00	.68	.84	.47
3	.95	.58	.76	.50	23	1.00	.16	.58	1.00
4	1.00	0.00	.50	1.00	24	1.00	.16	.58	1.00
5	1.00	.11	.55	1.00	25	.79	.16	.47	.78
6	1.00	.11	.55	1.00	26	1.00	.42	.71	.76
7	.95	.21	.58	.92	27	1.00	.37	.68	.81
8	.95	.21	.58	.92	28	.95	.42	.68	.68
9	.95	.11	.53	1.00	29	.89	.16	.53	.91
10	1.00	.58	.79	.58	30	1.00	.05	.53	1.00
11	1.00	.11	.55	1.00	31	1.00	.84	.92	.28*
12	.84	.16	.50	.85	32	1.00	.68	.84	.47
13	.89	.21	.55	.85	33	.95	.16	.55	.98*
14	1.00	.16	.58	1.00	34	1.00	.16	.58	1.00
15	.89	.32	.61	.73	35	1.00	.47	.74	.70
16	1.00	.37	.68	.81	36	.68	.11	.39	.73*
17	1.00	.16	.58	1.00	37	1.00	.53	.76	.65*
18	.89	.11	.50	.98	38	.95	.53	.74	.56
19	1.00	.26	.63	.93	39	1.00	.26	.63	.93*
20	1.00	.11	.55	1.00	40	1.00	.26	.63	.93

* ข้อสอบที่ลดลงเป็นฉบับที่ 2

ตาราง 12 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ
เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 2)

ข้อที่	P	r	ข้อที่	P	r
1	.74	.70	19	.63	.93
2	.50	1.00	20	.55	1.00
3	.76	.50	21	.84	.47
4	.50	1.00	22	.84	.47
5	.55	1.00	23	.58	1.00
6	.55	1.00	24	.58	1.00
7	.58	.92	25	.47	.78
8	.58	.92	26	.71	.76
9	.53	1.00	27	.68	.81
10	.79	.58	28	.68	.68
11	.55	1.00	29	.53	.91
12	.50	.85	30	.53	1.00
13	.55	.85	31	.84	.47
14	.58	1.00	32	.58	1.00
15	.61	.73	33	.74	.70
16	.68	.81	34	.74	.56
17	.58	1.00	35	.63	.93
18	.50	.98			

* ข้อสอบที่ลดลงเป็นฉบับที่ 3

ตาราง 13 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบเมื่อ
 ดัดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 3)

ข้อที่	P	r	ข้อที่	P	r
1	.74	.70	16	.68	.81
2	.50	1.00	17	.58	1.00
3	.76	.50	18	.50	.98
4	.50	1.00	19	.63	.93
5	.55	1.00	20	.55	1.00
6	.55	1.00	21	.84	.47*
7	.58	.92	22	.84	.47
8	.58	.92	23	.58	1.00
9	.53	1.00	24	.58	1.00*
10	.79	.58	25	.47	.78
11	.55	1.00	26	.71	.76*
12	.50	.85	27	.68	.81
13	.55	.85	28	.68	.68*
14	.58	1.00	29	.53	.91*
15	.61	.73	30	.53	1.00

* ข้อสอบที่ลดลงเป็นฉบับที่ 4

ตาราง 14 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบเมื่อ
ลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 4)

ข้อที่	P	r	ข้อที่	P	r
1	.74	.70	14	.58	1.00
2	.50	1.00	15	.61	.73
3	.76	.50	16	.68	.81
4	.50	1.00	17	.58	1.00
5	.55	1.00	18	.50	.98
6	.55	1.00	19	.63	.93
7	.58	.92	20	.55	1.00
8	.58	.92	21	.84	.47*
9	.53	1.00	22	.58	1.00*
10	.79	.58	23	.47	.78*
11	.55	1.00	24	.68	.81*
12	.50	.85	25	.53	1.00*
13	.55	.85			

* ข้อสอบที่ลดลงเป็นฉบับที่ 5

ตาราง 15 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ
เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 5)

ข้อที่	P	r	ข้อที่	P	r
1	.74	.70	11	.55	1.00*
2	.50	1.00	12	.50	.85*
3	.76	.50	13	.55	.85
4	.50	1.00	14	.58	1.00
5	.55	1.00	15	.61	.73
6	.55	1.00	16	.68	.81*
7	.58	.92	17	.58	1.00
8	.58	.92	18	.50	.98
9	.53	1.00	19	.63	.93
10	.79	.58*	20	.55	1.00*

* ข้อสอบที่ลดลงเป็นฉบับที่ 6

ตาราง 16 แสดงค่าความยาก (P) และค่าอำนาจจำแนก (r) ของแบบทดสอบเลือกตอบ
เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ (ฉบับที่ 6)

ข้อที่	P	r	ข้อที่	P	r
1	.74	.70	9	.53	1.00
2	.50	1.00	10	.55	.85
3	.76	.50	11	.58	1.00
4	.50	1.00	12	.61	.73
5	.55	1.00	13	.58	1.00
6	.55	1.00	14	.50	.98
7	.58	.92	15	.63	.93
8	.58	.92			

ภาคผนวก ข

**การคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ
เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู**

ตาราง 17 แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์จุดพิสัยของครู (ฉบับที่ 7)

จุดประสงค์ เชิงพฤติกรรมข้อที่	ข้อสอบ ข้อที่	อาจารย์ผู้สอนคนที่					ΣR	ค่าดัชนีความสอดคล้อง (IOC)
		1	2	3	4	5		
1	1	1	1	1	1	1	5	1.00
	2	1	1	1	1	1	5	1.00
	3	1	1	1	1	1	5	1.00
	10	1	1	1	1	-1	3	0.60
	21	1	1	1	-1	1	3	0.60
	22	1	1	1	1	1	5	1.00
	31	-1	-1	-1	1	1	-1	-0.20 **
	32	1	1	1	1	1	5	1.00
2	4	1	1	1	1	1	5	1.00
	5	1	1	1	1	1	5	1.00
	6	1	1	1	1	1	5	1.00
	11	1	1	1	-1	1	3	0.60
	23	-1	1	-1	1	1	1	0.20 **
	24	1	1	1	1	1	5	1.00
	33	1	-1	1	1	1	3	0.60
	34	1	1	1	1	-1	3	0.60
3	7	1	1	1	1	1	5	1.00
	8	1	1	1	1	1	5	1.00
	9	1	1	1	1	1	5	1.00
	12	1	1	1	1	1	5	1.00
	25	1	1	1	-1	-1	1	0.20 **
	26	-1	-1	1	1	1	1	0.20
	35	1	1	-1	1	1	3	0.60
	36	1	1	1	1	1	5	1.00
4	13	1	1	1	1	1	5	1.00
	14	1	1	1	1	1	5	1.00
	15	1	1	1	1	-1	3	0.60
	16	1	1	1	-1	1	3	0.60
	27	-1	1	-1	1	1	1	0.20 **
	28	1	-1	1	1	1	3	0.60
	37	1	1	1	1	1	5	1.00
	38	1	1	1	1	1	5	1.00
5	17	1	1	1	1	1	5	1.00
	18	1	1	1	1	-1	3	0.60
	19	1	1	1	-1	1	3	0.60
	20	1	1	1	1	1	5	1.00
	29	-1	-1	-1	1	1	-1	-0.20 **
	30	1	1	1	1	1	5	1.00
	39	1	1	1	1	1	5	1.00
	40	1	1	1	1	1	5	1.00

** ข้อสอบที่ลดลงจากฉบับที่ 1

ตาราง 18 แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบ
เมื่อลดจำนวนข้อ โดยใช้เกณฑ์คุณภาพของครู (ฉบับที่ 8)

จุดประสงค์ เชิงพฤติกรรมข้อที่	ข้อสอบ ข้อที่	อาจารย์ผู้สอนคนที่					ΣR	ค่าดัชนีความสอดคล้อง (IOC)
		1	2	3	4	5		
1	1	1	1	1	1	1	5	1.00
	2	1	1	1	-1	1	3	0.60
	3	1	1	1	1	1	5	1.00
	11	1	1	1	1	1	5	1.00
	21	1	1	1	1	-1	3	0.60
	22	1	1	-1	1	1	3	0.60
	31	-1	-1	1	1	1	1	0.20 **
2	4	1	1	1	-1	1	3	0.60
	5	1	1	1	1	-1	3	0.60
	6	1	1	1	1	1	5	1.00
	12	1	1	1	1	1	5	1.00
	23	1	-1	1	1	1	3	0.60
	24	1	1	1	1	1	5	1.00
	32	-1	1	-1	1	1	1	0.20 **
3	7	1	1	1	1	1	5	1.00
	8	-1	1	-1	1	1	1	0.20
	9	1	-1	1	1	1	3	0.60
	13	1	1	1	1	1	5	1.00
	25	1	1	1	1	1	5	1.00
	26	1	1	1	1	1	5	1.00
	33	1	1	1	-1	-1	1	0.20 **
4	10	1	1	1	1	1	5	1.00
	14	1	-1	1	1	1	3	0.60
	15	1	1	1	1	-1	3	0.60
	16	1	1	1	1	1	5	1.00
	27	-1	1	1	1	1	3	0.60
	28	1	1	1	1	1	5	1.00
	34	1	1	-1	-1	1	1	0.20 **
5	17	1	1	1	1	-1	3	0.60
	18	1	1	1	1	1	5	1.00
	19	1	1	1	1	1	5	1.00
	20	1	1	1	-1	1	3	0.60
	29	1	1	-1	1	1	3	0.60
	30	1	-1	1	1	1	3	0.60
	35	-1	1	1	1	1	3	0.60 **

** ข้อสอบที่ลดลงจากฉบับที่ 7

ตาราง 19 แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์คุณภาพของครู (ฉบับที่ 9)

จุดประสงค์ เชิงพฤติกรรมข้อที่	ข้อสอบ ข้อที่	อาจารย์ผู้สอนคนที่					ΣR	ค่าดัชนีความสอดคล้อง (IOC)
		1	2	3	4	5		
1	1	1	1	1	1	-1	3	0.60
	2	1	1	1	1	1	5	1.00
	3	1	1	1	1	1	5	1.00
	11	1	1	-1	1	1	3	0.60
	21	-1	-1	1	-1	1	-1	-0.20 **
	22	1	1	1	1	1	5	1.00
2	4	1	1	1	-1	1	3	0.60
	5	1	1	1	1	-1	3	0.60
	6	1	1	1	1	1	5	1.00
	12	1	-1	1	1	1	3	0.60
	23	-1	1	-1	1	1	1	0.20 **
	24	1	1	1	1	1	5	1.00
3	7	1	1	1	1	1	5	1.00
	8	1	1	1	1	1	5	1.00
	9	1	1	1	1	1	5	1.00
	13	1	1	-1	-1	1	1	0.20
	25	1	1	1	1	1	5	1.00
	26	-1	-1	1	1	-1	-1	-0.20 **
4	10	-1	1	1	1	1	3	0.60
	14	1	1	1	-1	1	3	0.60
	15	1	1	1	1	1	5	1.00
	16	1	1	1	1	1	5	1.00
	27	1	-1	-1	1	1	1	0.20 **
	28	1	1	1	1	-1	3	0.60
5	17	-1	1	1	1	1	3	0.60
	18	1	-1	1	1	1	3	0.60
	19	1	1	1	1	1	5	1.00
	20	1	1	1	1	1	5	1.00
	29	1	1	1	1	1	5	1.00
	30	1	1	-1	-1	-1	-1	-0.20 **

** ข้อสอบที่ลดลงจากฉบับที่ 8

ตาราง 20 แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 10)

จุดประสงค์ เชิงพฤติกรรมข้อที่	ข้อสอบ ข้อที่	อาจารย์ผู้สอนคนที่					ΣR	ค่าดัชนีความสอดคล้อง (IOC)
		1	2	3	4	5		
1	1	1	1	1	1	1	5	1.00
	2	-1	1	1	1	1	3	0.60
	3	1	1	1	1	-1	3	0.60
	11	1	1	1	1	1	5	1.00
	21	1	-1	-1	-1	1	-1	-0.20 **
2	4	1	1	1	1	1	5	1.00
	5	1	1	1	1	1	5	1.00
	6	1	1	-1	1	1	3	0.60
	12	1	1	1	1	1	5	1.00
	22	-1	-1	1	-1	-1	-3	-0.60 **
3	7	1	1	1	1	1	5	1.00
	8	1	1	1	1	-1	3	0.60
	9	1	1	1	1	1	5	1.00
	13	1	1	1	1	1	5	1.00
	23	-1	-1	-1	-1	1	-3	-0.60 **
4	10	1	1	1	1	1	5	1.00
	14	1	1	1	1	1	5	1.00
	15	1	1	1	1	1	5	1.00
	16	1	-1	1	1	1	3	0.60
	24	-1	1	-1	-1	-1	-3	-0.60 **
5	17	1	1	1	1	1	5	1.00
	18	1	1	1	1	1	5	1.00
	19	1	-1	1	1	1	3	0.60
	20	1	1	-1	1	1	3	0.60
	25	-1	1	1	-1	-1	-1	-0.20 **

** ข้อสอบที่ลดลงจากฉบับที่ 9

ตาราง 21 แสดงการคำนวณค่าดัชนีความสอดคล้อง (IOC) ของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของครู (ฉบับที่ 11)

จุดประสงค์ เชิงพฤติกรรมข้อที่	ข้อสอบ ข้อที่	อาจารย์ผู้สอนคนที่					ΣR	ค่าดัชนีความสอดคล้อง (IOC)
		1	2	3	4	5		
1	1	1	1	1	1	1	5	1.00
	2	1	1	1	1	1	5	1.00
	3	1	1	-1	1	1	3	0.60
	11	-1	-1	1	-1	-1	-3	-0.60 **
2	4	-1	1	1	1	1	3	0.60
	5	1	1	1	1	1	5	1.00
	6	1	1	1	1	1	5	1.00
	12	1	1	-1	-1	-1	-1	-0.20 **
3	7	1	1	1	1	1	5	1.00
	8	1	1	1	1	1	5	1.00
	9	1	1	1	1	1	5	1.00
	13	-1	-1	-1	-1	-1	-5	-1.00 **
4	10	1	-1	1	1	1	3	0.60
	14	1	1	1	1	1	5	1.00
	15	1	1	1	1	1	5	1.00
	16	-1	1	-1	-1	-1	-3	-0.60 **
5	17	1	1	1	1	1	5	1.00
	18	-1	-1	1	-1	1	-1	-0.20 **
	19	1	1	-1	1	1	3	0.60
	20	1	1	1	1	-1	3	0.60

** ข้อสอบที่ลดลงจากฉบับที่ 10

ภาคผนวก ค
แบบทดสอบที่ใช้ในการวิจัย

แบบทดสอบวิชาคณิตศาสตร์ (ฉบับที่ 1 ทักษะในการแก้โจทย์ปัญหา)
 ชั้นประถมศึกษาปีที่ 5 เรื่อง ทศนิยม เศษส่วน และโจทย์ปัญหาร้อยละ

คำชี้แจง

1. แบบทดสอบฉบับนี้มีทั้งหมด 40 ข้อ ใช้เวลา 80 นาที
2. แบบทดสอบเป็นแบบปรนัยชนิดเลือกตอบทั้งสิ้น กล่าวคือในแต่ละข้อ จะประกอบด้วยคำถามหรือปัญหา ซึ่งเป็นแบบประโยคภาษาและอีกส่วนหนึ่งเป็นคำตอบ ซึ่งมี 4 ตัวเลือก คือ ก. ข. ค. ง.
3. ให้นักเรียนพิจารณาตัวคำถามพร้อมกับคำตอบทั้ง 4 ตัวเลือก ว่าตัวเลือกใดเป็นคำตอบที่ถูกต้องที่สุด ซึ่งในแต่ละข้อจะมีคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียว เมื่อนักเรียนเลือกคำตอบที่ถูกต้องได้แล้ว ให้ทำเครื่องหมายกากบาท (X) ตรงตัวเลือกที่ถูกในกระดาษคำตอบ เช่น ถ้าข้อ ก. เป็นข้อถูก ให้ขีดกากบาท ดังตัวอย่าง

ข้อ ก. ข. ค. ง.

(0)

		X	
--	--	---	--

4. ถ้านักเรียนขีดคำตอบไปแล้ว แต่ต้องการเปลี่ยนคำตอบใหม่ เช่น ต้องการเปลี่ยนจากข้อ ก. เป็นข้อ ค. ก็ให้ขีดกากบาททับข้อ ก. แล้วขีดเส้นทับข้อ ค. ดังตัวอย่าง

ข้อ ก. ข. ค. ง.

(0)

X		X	
---	--	---	--

5. ในการขีดคำตอบแต่ละข้อ จะต้องขีดคำตอบเพียงช่องเดียวเท่านั้น ถ้าขีดเกินหนึ่งช่องจะถือว่าทำข้อนั้นผิด
6. ถ้าพบข้อใดยาก ก็ให้เว้นไปทำข้ออื่นต่อไปก่อน เมื่อมีเวลาเหลือจึงค่อยย้อนกลับมาทำข้อเดิมใหม่ เพราะอาจจะมีข้อที่นักเรียนทำได้อยู่ในข้อหลัง ๆ ก็ได้ และนักเรียนจงพยายามทำให้ครบทุกข้อด้วย

1. สุดามีเงินจำนวนหนึ่งแล้วซื้อสมุดไป $\frac{3}{5}$ บาท จ่ายค่าขนม $\frac{1}{5}$ บาท สุดาใช้เงินไปเท่าไร? ก. $\frac{1}{5}$ บาท ข. $\frac{2}{5}$ บาท ค. $\frac{1}{5}$ บาท ง. $\frac{4}{5}$ บาท	5. ถังใบหนึ่งมีน้ำอยู่ $\frac{10}{12}$ ลูกบาศก์เมตร ใช้ไป แล้ว $\frac{3}{4}$ ลูกบาศก์เมตร อยากทราบว่ายังเหลือ น้ำเท่าไร ? ก. $\frac{1}{12}$ ลูกบาศก์เมตร ข. $\frac{5}{12}$ ลูกบาศก์เมตร ค. $\frac{7}{12}$ ลูกบาศก์เมตร ง. $\frac{9}{12}$ ลูกบาศก์เมตร
2. สมศรีมีผ้าตัดเสื้ออยู่ $\frac{5}{6}$ เมตร สมใจมี $\frac{2}{6}$ เมตร สมศรีมีผ้ามากกว่าสมใจเท่าไร? ก. $\frac{1}{2}$ เมตร ข. $\frac{1}{6}$ เมตร ค. $\frac{4}{6}$ เมตร ง. 1 เมตร	6. อ้อยมีผ้าตัดเสื้ออยู่ $4\frac{5}{6}$ เมตร นำไปตัดเสื้อ $1\frac{3}{5}$ เมตร อ้อยยังเหลือผ้าตัดเสื้ออยู่เท่าไร? ก. $3\frac{2}{30}$ เมตร ข. $3\frac{7}{30}$ เมตร ค. $3\frac{8}{30}$ เมตร ง. $5\frac{7}{30}$ เมตร
3. บุญนำจ่ายค่าอาหารเป็นเงิน $\frac{5}{11}$ บาท เป็นค่า เช่าบ้าน $\frac{2}{11}$ บาท บุญนำใช้เงินไปคิดเป็นเท่าไร? ก. $\frac{3}{11}$ บาท ข. $\frac{7}{11}$ บาท ค. $\frac{3}{22}$ บาท ง. $\frac{7}{22}$ บาท	7. ปริชามีมะม่วง $\frac{3}{4}$ กิโลกรัม สมพรมีเป็น 3 เท่าของปริชา สมพรมีมะม่วงเท่าไร? ก. $\frac{3}{7}$ กิโลกรัม ข. $\frac{3}{12}$ กิโลกรัม ค. $2\frac{1}{4}$ กิโลกรัม ง. $3\frac{1}{4}$ กิโลกรัม
4. ลวดเส้นหนึ่งยาว $\frac{1}{4}$ เมตร ลวดเส้นที่สองยาว $\frac{9}{16}$ เมตร ลวดสองเส้นรวมกันยาวกี่เมตร? ก. $\frac{5}{16}$ เมตร ข. $\frac{8}{16}$ เมตร ค. $\frac{10}{16}$ เมตร ง. $\frac{13}{16}$ เมตร	8. ชั้นประถมปีที่ 5 มีนักเรียน 48 คน ขาด เรียน $\frac{1}{24}$ ของนักเรียนทั้งหมด อยากทราบว่า นักเรียน มาเรียนกี่คน? ก. 42 คน ข. 44 คน ค. 46 คน ง. 47 คน

9. มีน้ำตาลทรายอยู่ $5\frac{1}{4}$ กิโลกรัม แบ่งใส่ถุง ๆ ละ $\frac{3}{4}$ กิโลกรัม จะแบ่งใส่ถุงได้ทั้งหมดกี่ถุง? ก. 2 ถุง ข. 4 ถุง ค. 7 ถุง ง. 8 ถุง	13. ถ้านิพนธ์มีน้ำหนักตัวเพิ่มขึ้นอีก 3.85 กิโลกรัม ปัจจุบันเขาชั่งได้ 57.65 กิโลกรัม เดิมนิพนธ์มีน้ำหนักตัวเท่าไร? ก. 53.80 กิโลกรัม ข. 54.20 กิโลกรัม ค. 54.80 กิโลกรัม ง. 61.50 กิโลกรัม
10. โอ่งใบหนึ่งมีน้ำอยู่คิดเป็น $\frac{5}{8}$ ของโอ่ง ถ้าต้องการเติมน้ำให้เต็มโอ่ง จะต้องเติมน้ำลงไปอีกเท่าไร? ก. $\frac{1}{8}$ ข. $\frac{3}{8}$ ค. $\frac{6}{8}$ ง. $\frac{7}{8}$	14. ฝรั่งลูกหนึ่งหนัก 0.8 กิโลกรัม ส้มลูกหนึ่งหนัก 0.25 กิโลกรัม ฝรั่งหนักกว่าส้มเท่าไร? ก. 0.35 กิโลกรัม ข. 0.55 กิโลกรัม ค. 4.5 กิโลกรัม ง. 5.6 กิโลกรัม
11. มีแตงโมผลหนึ่งแบ่งให้เพื่อนไปแล้ว $\frac{3}{8}$ คงเหลือแตงโมคิดเป็นเศษส่วนเท่าไร? ก. $\frac{1}{8}$ ข. $\frac{3}{8}$ ค. $\frac{5}{8}$ ง. $\frac{7}{8}$	15. พยางค์กางเกงตัวละ 359.75 บาท ซื้อเสื้อ 2 ตัว ตัวแรกราคา 75.50 บาท ตัวที่สองราคา 108.50 บาท เขาจ่ายเงินซื้อกางเกงมากกว่าซื้อเสื้อเท่าไร? ก. 175.75 บาท ข. 176.75 บาท ค. 284.25 บาท ง. 543.75 บาท
12. ทศพลมีเงิน 80 บาท ใช้จ่ายไป $\frac{2}{5}$ ของเงินที่มีอยู่ ทศพลเหลือเงินกี่บาท? ก. 16 บาท ข. 32 บาท ค. 48 บาท ง. 52 บาท	16. สุนัขสูง 145 เซนติเมตร สุภาสูง 1 เมตร 27 เซนติเมตร สุนัขสูงกว่าสุภาเท่าไร? ก. 0.18 เมตร ข. 0.72 เมตร ค. 2.18 เมตร ง. 2.72 เมตร

17. ลัตดาสอบได้ 420 คะแนน จากคะแนนเต็ม 500 คะแนน ลัตดาสอบได้กี่เปอร์เซ็นต์? ก. 80% ข. 81% ค. 82% ง. 84%	21. เมตตามีเชือกเส้นหนึ่งยาว $\frac{4}{15}$ เมตร สมชายมีเชือกน้อยกว่าเมตดาอยู่ $\frac{1}{15}$ เมตร สมชายมีเชือกยาวเท่าไร? ก. $\frac{1}{15}$ เมตร ข. $\frac{2}{15}$ เมตร ค. $\frac{3}{15}$ เมตร ง. $\frac{5}{15}$ เมตร
18. ซื้อพัดลมมาราคา 640 บาท ต้องการขายได้กำไร 5% จะต้องขายในราคาเท่าไร? ก. 608 บาท ข. 672 บาท ค. 675 บาท ง. 680 บาท	22. สมพิศมีลวดยาว $5\frac{7}{8}$ เมตร สมบัติมีลวดยาว $3\frac{4}{8}$ เมตร สมบัติมีลวดน้อยกว่าสมพิศกี่เมตร? ก. $2\frac{3}{8}$ เมตร ข. $2\frac{4}{8}$ เมตร ค. $3\frac{3}{8}$ เมตร ง. $8\frac{3}{8}$ เมตร
19. ชาวนาขายข้าวได้เงิน 7,500 บาท เก็บไว้ทำทุน 60% จงหาว่าชาวนาเก็บเงินไว้ทำทุนเท่าไร? ก. 3,500 บาท ข. 4,000 บาท ค. 4,200 บาท ง. 4,500 บาท	23. มีแก๊สอยู่ในถัง $\frac{8}{10}$ ลูกบาศก์เมตร ใช้ไป $\frac{1}{2}$ ลูกบาศก์เมตร จะเหลือแก๊สในถังเท่าไร? ก. $\frac{3}{10}$ ลูกบาศก์เมตร ข. $\frac{4}{10}$ ลูกบาศก์เมตร ค. $\frac{7}{10}$ ลูกบาศก์เมตร ง. $\frac{8}{10}$ ลูกบาศก์เมตร
20. ลงทุนซื้อของมา 200 บาท ต้องการขายได้กำไร 6% จะต้องขายในราคาเท่าไร? ก. 206 บาท ข. 212 บาท ค. 260 บาท ง. 306 บาท	24. มะลิซื้อฝรั่งหนัก $2\frac{2}{3}$ กิโลกรัม ซื้อมะม่วงหนัก $3\frac{1}{2}$ กิโลกรัม มะลิซื้อผลไม้ทั้งหมดกี่กิโลกรัม? ก. $5\frac{1}{6}$ กิโลกรัม ข. $5\frac{3}{5}$ กิโลกรัม ค. $6\frac{1}{6}$ กิโลกรัม ง. $6\frac{1}{3}$ กิโลกรัม

25. ที่ดินมีพื้นที่ 12 ไร่ ต้องการแบ่งเป็นแปลง ๆ ละ $1\frac{1}{3}$ ไร่ จะแบ่งได้กี่แปลง? ก. 6 แปลง ข. 8 แปลง ค. 9 แปลง ง. 16 แปลง	29. มะม่วง 450 ผล ราคา 630 บาท อยากทราบว่ามะม่วงราคาร้อยละเท่าไร? ก. 140 บาท ข. 240 บาท ค. 280 บาท ง. 360 บาท
26. มีเงินอยู่ 150 บาท แบ่งให้น้อง $\frac{7}{30}$ ของเงินที่มีอยู่ น้องจะได้เงินเท่าไร? ก. 35 บาท ข. 45 บาท ค. 55 บาท ง. 65 บาท	30. ในการสอบครั้งหนึ่งมีคะแนนเต็ม 500 คะแนน สมปองสอบได้ 480 คะแนน คิดเป็นกี่เปอร์เซ็นต์ของคะแนนเต็ม? ก. 85 % ข. 88 % ค. 90 % ง. 96 %
27. แดงกระโดดได้ไกล 5.40 เมตร ดำกระโดดได้ 6.15 เมตร ดำกระโดดได้ไกลกว่าแดงเท่าไร? ก. 0.75 เมตร ข. 0.85 เมตร ค. 1.25 เมตร ง. 1.65 เมตร	31. ตาสีมีวัว $\frac{7}{7}$ ตัว ขายไป $\frac{4}{7}$ ตัว ตาสียังมีวัวอยู่อีกเท่าไร? ก. $\frac{2}{7}$ ตัว ข. $\frac{3}{7}$ ตัว ค. $\frac{5}{7}$ ตัว ง. $\frac{7}{7}$ ตัว
28. เอนกสอยมะม่วงได้ 32.35 กิโลกรัม อนันต์สอยได้ 29.2 กิโลกรัม ทั้งสองสอยมะม่วงรวมกันเป็นเท่าไร? ก. 3.15 กิโลกรัม ข. 60.55 กิโลกรัม ค. 61.37 กิโลกรัม ง. 61.55 กิโลกรัม	32. ไก่ตัวหนึ่งซึ่งได้ $\frac{15}{10}$ กิโลกรัม เปิดตัวหนึ่งซึ่งได้ $\frac{3}{10}$ กิโลกรัม ไก่และเปิดตัวซึ่งพร้อมกันจะหนักเท่าไร? ก. $\frac{12}{10}$ กิโลกรัม ข. $\frac{18}{10}$ กิโลกรัม ค. $\frac{19}{10}$ กิโลกรัม ง. $\frac{18}{20}$ กิโลกรัม

33. มาลีมีเชือกยาว $\frac{2}{5}$ เมตร วิภามีเชือกยาว $\frac{3}{10}$ เมตร รวมสองคนมีเชือกยาวเท่าไร? ก. $\frac{5}{5}$ เมตร ข. $\frac{5}{10}$ เมตร ค. $\frac{6}{10}$ เมตร ง. $\frac{7}{10}$ เมตร	37. เอกเติมน้ำมันรถยนต์เมื่อตอนเช้า 22.5 ลิตร ตอนเย็นเติมน้ำมันอีก 37.5 ลิตร วันนี้เอกเติมน้ำมันทั้งหมดกี่ลิตร? ก. 58.5 ลิตร ข. 59.0 ลิตร ค. 60.0 ลิตร ง. 61.5 ลิตร
34. ช่างทาสีเปิดกระป๋องสีพบว่าสีอยู่ $\frac{5}{6}$ ของกระป๋อง เขานำไปทาพื้นฝาบ้าน $\frac{2}{3}$ ของกระป๋องอยากทราบว่าสียังเหลือเท่าไร? ก. $\frac{1}{6}$ ของกระป๋อง ข. $\frac{2}{6}$ ของกระป๋อง ค. $\frac{3}{6}$ ของกระป๋อง ง. $\frac{4}{6}$ ของกระป๋อง	38. พรชัยซื้อหนังสือ 1 เล่ม ราคา 17.75 บาท สมุด 2 เล่ม ราคา 12.50 บาท และรองเท้า 1 คู่ ราคา 59 บาท พรชัยจะต้องจ่ายเงินทั้งหมดเท่าไร? ก. 30.25 บาท ข. 72.00 บาท ค. 79.75 บาท ง. 89.25 บาท
35. ถั่วเขียวถุงหนึ่งหนัก $\frac{1}{2}$ กิโลกรัม ถั่วเขียว 50 ถุงจะหนักเท่าไร? ก. 20 กิโลกรัม ข. 25 กิโลกรัม ค. 40 กิโลกรัม ง. 50 กิโลกรัม	39. แดงซื้อส้มมา 280 บาท แล้วขายไปได้กำไร 10% แดงขายส้มไปเป็นเงินเท่าไร? ก. 308 บาท ข. 309 บาท ค. 310 บาท ง. 311 บาท
36. มีที่ดินอยู่ 120 ไร่ แบ่งขายเป็นแปลง ๆ ละ $\frac{5}{6}$ ไร่ จะแบ่งที่ดินได้ทั้งหมดกี่แปลง? ก. 100 แปลง ข. 124 แปลง ค. 140 แปลง ง. 144 แปลง	40. สุรินทร์สอบได้ 85% จากคะแนนเต็ม 500 คะแนน สุรินทร์สอบได้คะแนนเท่าไร? ก. 315 คะแนน ข. 325 คะแนน ค. 425 คะแนน ง. 435 คะแนน

แบบทดสอบฉบับที่ 1 ที่คัดเลือกไว้จำนวน 40 ข้อ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ค่าสถิติ

ฉบับที่ 2 จำนวน 35 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 23 24
25 26 27 28 29 30 32 34 35 38 40

ฉบับที่ 3 จำนวน 30 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24
25 26 27 28 29 30

ฉบับที่ 4 จำนวน 25 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 22 23 25 27
30

ฉบับที่ 5 จำนวน 20 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20

ฉบับที่ 6 จำนวน 15 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 13 14 15 17
18 19

แบบทดสอบฉบับที่ 1 ที่คัดเลือกไว้จำนวน 40 ข้อ เมื่อลดจำนวนข้อ โดยใช้เกณฑ์ดุลพินิจของ
ครู

ฉบับที่ 7 จำนวน 35 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 24 26
28 30 32 33 34 35 36 37 38 39 40

ฉบับที่ 8 จำนวน 30 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24
25 26 27 28 29 30

ฉบับที่ 9 จำนวน 25 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 22 24 25 28
29

ฉบับที่ 10 จำนวน 20 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20

ฉบับที่ 11 จำนวน 15 ข้อ ได้แก่ ข้อ 1 2 3 4 5 6 7 8 9 10 14 15 17
19 20

ภาคผนวก ง

1. รายชื่อผู้เชี่ยวชาญที่ประเมินความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม
2. รายชื่ออาจารย์ผู้สอนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่พิจารณาคัดเลือกข้อสอบให้มีจำนวนข้อลดลง โดยใช้เกณฑ์ดุลพินิจของครู

รายชื่อ ผู้เชี่ยวชาญที่ประเมินความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม

อาจารย์บุพา	มานะจิตต์	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มศว. ประสานมิตร
อาจารย์อรนุช	ธิดิรัถย์พณิชย	อาจารย์โรงเรียนสวนกุหลาบวิทยาลัย
อาจารย์ชลธิชา	สีลวัตะ	อาจารย์โรงเรียนวัดราชบพิธ
อาจารย์จุรีรัตน์	นันทย์ทวีกุล	อาจารย์โรงเรียนวัดราชบพิธ
อาจารย์มลิวัลย์	รวยลาภ	อาจารย์โรงเรียนปากข้าวสาร จ. สระบุรี

รายชื่อ อาจารย์ผู้สอนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ ๕ ที่พิจารณาคัดเลือกข้อสอบให้มีจำนวนข้อลดลง โดยใช้เกณฑ์ดุลพินิจของครู

อาจารย์จิตบรรจง	น้อยสกุล	อาจารย์โรงเรียนบางปลาม้า จ. สุพรรณบุรี
อาจารย์ลลอง	การสาถิ	อาจารย์โรงเรียนวัดช่องลม จ. สุพรรณบุรี
อาจารย์นันทนา	จวิรัชวัชรกุล	อาจารย์โรงเรียนวัดอกรักษ์ จ. สุพรรณบุรี
อาจารย์สมควร	เข้มเกษม	อาจารย์โรงเรียนวัดอกรักษ์ จ. สุพรรณบุรี
อาจารย์จุฬารัตน์	โพธิ์ประดิษฐ์	อาจารย์โรงเรียนบรรหารแจ่มใสวิทยา 2 จ. สุพรรณบุรี

ประวัติย่อของผู้วิจัย

ชื่อ นางสาวพิมพ์อุษา เจริญยศ
 เกิดวันที่ 16 กรกฎาคม 2500
 สถานที่เกิด อำเภอบางปลาม้า จังหวัดสุพรรณบุรี
 สถานที่อยู่ปัจจุบัน 195 หมู่ 6 ตำบลโคกกราม อำเภอบางปลาม้า
 จังหวัดสุพรรณบุรี 72150
 สถานที่ทำงานปัจจุบันโรงเรียนวัดอรัญญิก ตำบลอรัญญิก อำเภอบางปลาม้า
 จังหวัดสุพรรณบุรี 72150
 ตำแหน่ง อาจารย์ 2 ระดับ 6

ประวัติการศึกษา

พ.ศ. 2517	มัธยมศึกษาปีที่ 3 จากโรงเรียนบางปลาม้า “สูงสมารผดุงวิทย์” จังหวัดสุพรรณบุรี
พ.ศ. 2522	ป.ศ.สูง (วิชาเอกคณิตศาสตร์) จากวิทยาลัยครูนครปฐม
พ.ศ. 2526	ศษ.บ. (วิชาเอกสังคมศึกษา) จากมหาวิทยาลัยสุโขทัยธรรมมาธิราช
พ.ศ. 2538	กศ.ม. (การวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบ เมื่อลดจำนวนข้อ
โดยใช้เกณฑ์ค่าสถิติกับใช้เกณฑ์ดุลพินิจของครู

บทคัดย่อ
ของ
พิมพ์ธูษา เจริญยศ

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
กุมภาพันธ์ 2539

การวิจัยนี้ เป็นการศึกษาการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเลือกตอบเมื่อลดจำนวนข้อโดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู โดยมีจุดมุ่งหมายดังนี้ 1) เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติ 2) เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครู 3) เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน โดยใช้แบบทดสอบวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหา ชั้นประถมศึกษาปีที่ 5 จำนวน 11 ฉบับ ทดสอบกับกลุ่มตัวอย่างจำนวน 608 คน แล้วคำนวณค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตรของเฟลด์ต์ ทดสอบความแตกต่างของ ค่าความเชื่อมั่นโดยใช้สูตรของมาราสกุโยโล และทดสอบความแตกต่างรายคู่ โดยใช้สูตรของเซฟเฟ่

ผลการศึกษาพบว่า 1) ค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ค่าสถิติพิจารณา แบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด และจากการเปรียบเทียบค่าความเชื่อมั่นแต่ละฉบับ พบว่า แบบทดสอบที่มีจำนวน 35 และ 15 ข้อ แตกต่างไปจากแบบทดสอบที่มีจำนวน 40 ข้อ ส่วนที่เหลือไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ 2) ค่าความเชื่อมั่นของแบบทดสอบระหว่างฉบับที่มีจำนวน 40 ข้อ กับฉบับที่ลดจำนวนข้อลงเหลือ 35, 30, 25, 20 และ 15 ข้อ โดยใช้เกณฑ์ดุลพินิจของครูพิจารณา แบบทดสอบที่มีจำนวน 25 ข้อ มีค่าความเชื่อมั่นสูงสุด ส่วนแบบทดสอบที่มีจำนวน 15 ข้อ มีค่าความเชื่อมั่นต่ำสุด และจากการเปรียบเทียบค่าความเชื่อมั่นแต่ละฉบับ พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติเป็นส่วนใหญ่ ยกเว้นค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวน 25 ข้อ เพียงฉบับเดียวเท่านั้นที่ให้ค่าความเชื่อมั่นไม่แตกต่างไปจากแบบทดสอบที่มีจำนวน 40 ข้อ 3) ค่าความเชื่อมั่นของแบบทดสอบที่ลดจำนวนข้อลง โดยใช้เกณฑ์ค่าสถิติ กับใช้เกณฑ์ดุลพินิจของครู ที่มีจำนวนข้อเท่ากัน พบว่า ค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทุกคู่ โดยสรุปแล้ว แบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ฉบับโจทย์ปัญหา ที่มีจำนวน 40 ข้อ สามารถลดจำนวนข้อลงเหลือเป็น 25 ข้อ ได้โดยยังคงมีค่าความเชื่อมั่นไม่เปลี่ยนแปลง

THE COMPARISON OF RELIABILITY OF THE MULTIPLE CHOICE TESTS WITH REDUCED
NUMBER OF ITEMS BY STATISTICAL AND TEACHER JUDGEMENT CRITERION

AN ABSTRACT

BY

PIMUSA CHAROENYOS

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

February 1996

The purposes of this study were three folds. First , it aimed to compare the reliability of multiple choice tests with 40 items and 35, 30, 25, 20 and 15 reduced number of items by statistical judgement criterion , second to compare the reliability of those by teacher judgement criterion , and third to compare the reliability of the test with reduced number of items in equal by statistical and teacher judgement criterion. Eleven Pratom Suksa 5's mathematical problem solving tests were administered to 608 subjects. The reliability was calculated by Feldt Fomular then the reliability difference of pair comparison were tested by Marascouilo and Scheffe⁴ Methods accordingly.

Major findings were as follows : 1) Among the reliability of multiple choice tests with 40 items and 35, 30, 25, 20 and 15 reduced number of items by statistical judgement criterion, it was found that the test with 25 items got the highest reliability while the test with 15 items got the lowest. When compared the reliability of each test, it was found that there was significant difference among the test with 35 and 15 items and 40 items while the other was no significant difference. 2) Among those by teacher judgement criterion, it was found that the test with 25 items got the highest reliability while the test with 15 items got the lowest. When compared the reliability of each test it was found that most of them was significant difference , except there was no significant difference among the test with 25 items and 40 items. 3) There was no significant difference among the reliability of reduced number of tests by statistical and teacher judgement criterion. Then it could concluded that mathematical problem solving test with 40 items could reduce to 25 items while there still was the same reliability.